THE ANNALS *of* MATHEMATICAL STATISTICS

THE ANNALS OF MATHEMATICAL STATISTICS IS AFFILIATED WITH THE AMERICAN STATISTICAL ASSOCIATION AND IS DEVOTED TO THE THEORY AND APPLICATION OF MATHEMATICAL STATISTICS

Volume VI, 1935

PUBLISHED QUARTERLY
ANN ARBOR, MICHIGAN

Four Dollars per annum

Made in United States of America

Address: ANNALS OF MATHEMATICAL STATISTICS
Post Office Box 171, Ann Arbor, Michigan

COMPOSED AND PRINTED AT THE
WAVERLY PRESS, INC.
BALTIMORE, MD.

SOME INTERESTING FEATURES OF FREQUENCY CURVES

By Richmond T. Zoch

Introduction

It is well known that in the normal error curve the points of inflection are equidistant from the mode. However it has never been pointed out that this is also a characteristic of all of the bell-shaped Pearson Frequency Curves. This fact can be most easily shown by placing the mode at the abscissa $x = 0$.

Many rough checks have been developed for use in applying the Theory of Least Squares. The second part of this paper develops a rough check on the computation for use when fitting a Pearson Frequency Curve to a set of observations. No rough checks on computation are given in textbooks on Pearson's Frequency Curves.

At present it is customary to follow a separate procedure for each Type of curve when computing the constants of a Pearson Frequency Curve. The third part of this paper shows how a single system may be followed for all Types. A single procedure is very desirable in order that the rough check of Part 2 may be quickly applied.

Part 1. Points of Inflection

Perhaps nothing brings out the limitations of the bell-shaped Pearson Curves in a more striking manner than a discussion of their points of inflection. In dealing with frequency curves it is well known that any curve can be fitted to a given distribution and that the real problem in curve fitting is the selection of a curve. Figures 1, 2, and 3 illustrate three hypothetical histograms. All three of these histograms are bell-shaped yet none of them will be closely fitted by any of the Pearson Curves. The reasons will be pointed out presently.

The differential equation from which Pearson derived his system of frequency curves is

$$\frac{dy}{dx} = \frac{y(x - P)}{b_2x^2 + b_1x + b_0}.$$

By putting $x - P = X$, i.e. by placing the mode at the abscissa $X = 0$, this differential equation may be written:

$$\frac{dy}{dX} = \frac{yX}{\pm B_2X \pm B_1X + B_0}$$

where the $+$ or $-$ sign is taken according to the type of the curve. (It will be shown later that the constant term of the denominator must be less than zero.)

Since in the Type III curve B_2 is 0 and in the "Normal Curve" both B_2 and B_1 are 0 it will be advantageous to consider the general case of

$$\frac{dy}{dX} = \frac{yX}{F(X)},$$

where $F(X)$ is an integral rational function of the n^{th} degree, at once rather than considering special cases first.

If

$$\frac{dy}{dX} = \frac{yX}{F(X)},$$

then

$$\frac{d^2y}{dX^2} = \frac{y}{[F(X)]^2}\{X^2 + F(X) - XF'(X)\}.$$

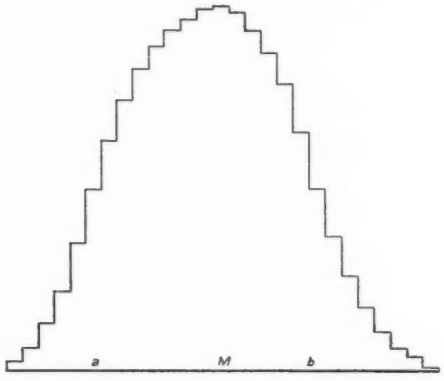

FIG. 1

In order to locate the points of inflection, $\frac{d^2y}{dX^2}$ is equated to zero. Then we have:

$$X^2 + F(X) - XF'(X) = 0. \tag{1}$$

This equation is always of the same degree as $F(X)$ except when $F(X)$ is linear or constant. Hence we have proved the Theorem: If $y = G(X)$ be the solution of the differential equation

$$\frac{dy}{dX} = \frac{yX}{F(X)},$$

then the number of points of inflection of y cannot exceed the degree of $F(X)$ when $F(X)$ is of degree greater than one.

Now $F(X) = B_nX^n + B_{n-1}X^{n-1} + \cdots + B_2X^2 + B_1X + B_0$. Whence equation (1) can be written in the form:

$$(1-n)B_nX^n + (2-n)B_{n-1}X^{n-1} + (3-n)B_{n-2}X^{n-2} + \cdots$$
$$+ (r-n)B_{n-r+1}X^{n-r+1} + \cdots - 3B_4X^4 - 2B_3X^3 + (1-B_2)X^2 + B_0 = 0.$$

Hence we have established the Theorem: The coefficient of the linear term of X in the equation of the points of inflection is zero.

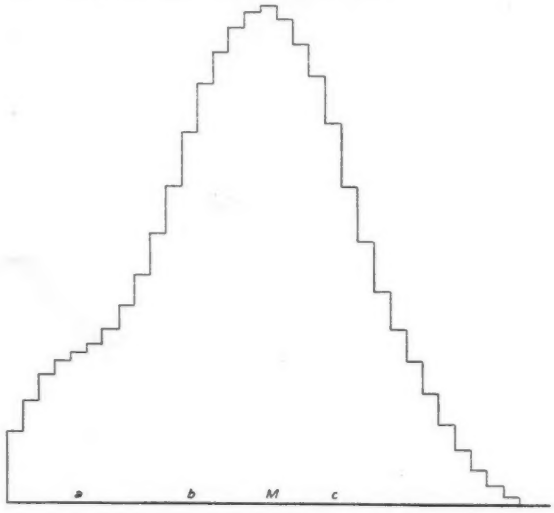

FIG. 2

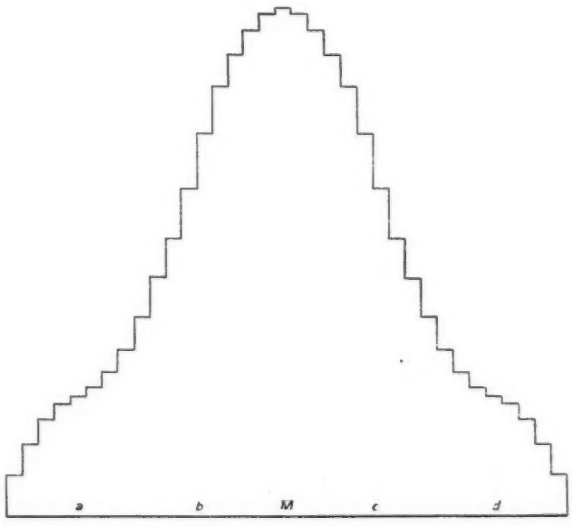

FIG. 3

For the "Normal Curve" and also for Type III,

$$B_2 = B_3 = B_4 = \cdots = B_n = 0.$$

Hence the points of inflection of these two Types are given by $X = \pm\sqrt{-B_0}$.

For Types I and II, B_2 is positive and $B_3 = B_4 = \cdots = B_n = 0$, and the

points of inflection are $X = \pm\sqrt{\frac{-B_0}{1 - B_2}}$. Hence the points of inflection are undefined if $B_2 = 1$, are pure imaginary if $B_2 > 1$, and real if $B_2 < 1$.

For Types IV, V, VI and VII, B_2 is negative and $B_3 = \cdots = B_n = 0$, and the points of inflection are at $X = \pm\sqrt{\frac{-B_0}{1 + |B_2|}}$.

In some of these Types it may happen that the abscissae of the points of inflection though real will lie beyond the range of the curve. Thus Types III and VI may have 1 or 2 points of inflection, the single point of inflection occurring when $\left|\sqrt{\frac{-B_0}{1 + B_2}}\right|$ > the range of the curve in the direction that the range is limited. Type II may have 0 or 2 points of inflection, as there will be no real points of inflection when $B_2 \geqq 1$. Type I may have 0, 1 or 2 points of inflection. Types IV, V and VII as well as the "Normal Curve" always have 2 and only two points of inflection.

Now it should be noted that when one of the eight bell-shaped Pearson curves has two points of inflection then the abscissae of these 2 points of inflection are equidistant from the abscissa of the mode. In figure 1 a point of inflection will be at abscissa b and another at abscissa a. (M is the abscissa of the mode.) Since $b - M \neq M - a$ none of the Pearson curves will fit this histogram closely. In figure 2, points of inflection occur at abscissae a, b, and c. Since a Pearson curve can have at most two points of inflection no Pearson curve will fit this histogram closely. In figure 3 there are four points of inflection and no Pearson curve will fit this histogram closely.

Part 2. Range

DEFINITION: A bell-shaped curve is a continuous curve which starts at zero (or zero as a limit), rises to a single maximum, at which maximum point the first derivative is zero, and then falls to zero (or zero as a limit).

Or, more formally, $y = G(x)$ is a bell-shaped curve if $G(x_1) = G(x_2) = 0$ and if $G'(P) = 0$ and $G''(P) < 0$ where $G(x)$ is continuous and does not vanish in the interval from x_1 to x_2 and P is a unique point in this interval.

If a bell-shaped curve has the value of zero at two finite points, one on each side of the maximum (mode), it is said to be of limited range in both directions, or briefly, of limited range.

If a bell-shaped curve has the value of zero at only one finite point it is said to be of limited range in one direction, or also of unlimited range in one direction.

If a bell-shaped curve has the value of zero only at $\pm\infty$, i.e. at no finite points, it is said to be of unlimited range in both directions, or briefly, of unlimited range.

THEOREM I: If $F(x)$ can be separated into a finite number of factors each either of the form $(x - r_i)$ or $(x^2 + 2r_j x + r_j^2 + r_{0_j}^2)$ where no real root is repeated and $y = G(x)$ is a *bell-shaped* curve which is a solution of the differential equation

$$\frac{dy}{dx} = \frac{y(x - P)}{F(x)},$$

then if $F(x)$ has no real roots, y is of unlimited range in both directions; if all of the real roots of $F(x)$ lie on the same side of P, y is of limited range in one (that) direction; if at least one real root of $F(x)$ lies on one side of P and at least one on the other side, y is of limited range in both directions.

PROOF: If $F(x) = 0$ when $x = P$, we have

$$\frac{dy}{dx} = \frac{y}{g(x)}$$

where $g(x) = F(x) \div (x - P)$. This derivative is zero only when $y = 0$ or $g(x) = \pm\infty$. Hence the solution does not have a finite maximum and therefore is not a bell-shaped curve. If $F(x) > 0$ when $x = P$, we have

$$\frac{d^2y}{dx^2} = \frac{y}{[F(x)]^2}\left[(x - P)^2 + F(x) - (x - P)\frac{d}{dx}F(x)\right]$$

and

$$\left.\frac{d^2y}{dx^2}\right|_{x=P} = \frac{y}{[F(x)]^2}[F(x)]$$

which is greater than zero and, since at a maximum the second derivative must not be greater than zero, in this case the solution would have a minimum at $x = P$ and therefore would not be a bell-shaped curve. As the theorem concerns only those solutions which are bell-shaped curves, $F(x) < 0$ when $x = P$. If $F(x) = 0$ when $x \neq P$ then $\frac{dy}{dx} = \pm\infty$ unless y is also zero. Assume $y \neq 0$. Since $F(x)$ is negative, if $y \neq 0$ when $F(x) = 0$ then $\frac{dy}{dx} \to -\infty$ as $F(x) \to 0$, for an $x > P$, and changes to $+\infty$ as $F(x)$ changes sign on passing through the value 0. Hence the curve would contain another maximum before falling to zero and therefore the solution is not a bell-shaped curve. Similar reasoning holds for an $x < P$. Therefore if $y \neq 0$ when $F(x) = 0$, the curve is not bell-shaped. If $y = 0$ when $F(x) = 0$, the curve has its range limited at this point. That is, any real number which makes $F(x)$ vanish will also make y vanish if y represents a bell-shaped curve. Hence if all of the real roots lie on the same side of P the curve is of limited range in that direction only, while if at least one of the real roots lies on each side of P the curve is of limited range in both directions. If $F(x)$ contains no real roots it does not vanish for any real value of x. In this case, by partial fractions the differential equation becomes:

$$\frac{dy}{y} = \frac{k_{11}\,dx}{(x + r_1)^2 + r_{0_1}^2} + \frac{k_{21}\,dx}{(x + r_2)^2 + r_{0_2}^2} + \cdots + \frac{2k_{21}(x + r_1)\,dx}{(x + r_1)^2 + r_{0_1}^2} + \frac{2k_{22}(x + r_2)\,dx}{(x + r_2)^2 + r_{0_2}^2} + \cdots.$$

On integrating,

$$y = C\left[(x+r_1)^2 + r_{0_1}\right]^{k_{21}}\left[(x+r_2)^2 + r_{0_2}^2\right]^{k_{22}} \cdots e^{k_{11} \arctan \frac{x+r_1}{r_{0_1}} + \cdots}.$$

Hence y does not vanish for a finite real value of x and the Theorem is fully established.

THEOREM II: If $F(x)$ can be separated into a finite number of factors each either of the from $(x - r_i)$ or $(x^2 + 2r_jx + r_j^2 + r_{0_j}^2)$ where no real root is repeated and $y = G(x)$ is a *bell-shaped* curve which is a solution of the differential equation $\frac{dy}{dx} = \frac{y(x-P)}{F(x)}$, then if y is of unlimited range, $F(x)$ contains no real roots; if y is of limited range in one direction, all of the real roots of $F(x)$ lie on the same (that) side of P; if y is of limited range in both directions, at least one of the real roots of $F(x)$ lies on one side of P and at least one on the other.

PROOF: By partial fractions the differential equation may be written:

$$\frac{dy}{y} = \frac{k_{11}\,dx}{x - r_{11}} + \frac{k_{21}\,dx}{x - r_{12}} + \cdots + \frac{k_{21}\,dx}{(x+r_{21})^2 + r_{0_1}^2} + \frac{k_{22}\,dx}{(x+r_{22})^2 + r_{0_2}^2} + \cdots \frac{2k_{31}(x+r_{21})\,dx}{(x+r_{21})^2 + r_{0_1}^2} + \frac{2k_{32}(x+r_{23})\,dx}{(x+r_{22})^2 + r_{0_2}^2} + \cdots$$

and on integrating:

$$y = C(x - r_{11})^{k_{11}}(x - r_{12})^{k_{12}} \cdots \left[(x+r_{21})^2 + r_{0_1}^2\right]^{k_{31}} \cdots e^{k_{21} \arctan \frac{x+r_{21}}{r_{0_1}} + \cdots}.$$

Hence $y = 0$ for $x = r_{11}, r_{12}, \cdots$ and for no other finite values of x provided k_{11}, $k_{12}, \cdots$ are positive. If one or more of the k_{ij} are negative, $y = \infty$ at such points and unless some r_{ij} closer to P has previously made y vanish, the curve is not bell-shaped. Therefore, for bell-shaped curves, the exponent of the factor containing the real root of smallest absolute value on each side of P is positive. Therefore: if y is of limited range in both directions, at least one real root lies on each side of P; if y is of unlimited range in one direction, all of the real roots lie on the same side of P; if y is of unlimited range it contains no real roots. Hence the Theorem is established.

The effect of repeated real roots will now be considered. If a real root is repeated an odd number of times at $x = r$, then $F(x)$ changes sign at $x = r$ and the first theorem is true. If a real root is repeated an even number of times at $x = r$, then $F(x)$ does not change sign at $x = r$ and we know that either (a) $y = 0$ at $x = r$; or (b) y is finite and $\neq 0$ and $\frac{dy}{dx} = \pm\infty$ at $x = r$, i.e. there is a point of inflection at $x = r$. It will now be shown that (b) cannot occur. If case (b) is possible, y is continuous at $x = r$, $\frac{dy}{dx} = \pm\infty$ according as $(r - P) \lesseqgtr 0$

moreover $\frac{dy}{dx}$ does not change sign in the neighborhood of the point $x = r$, and $\frac{d^2y}{dx^2}$ changes sign from $+\infty$ to $-\infty$ or vice versa according as $(r - P) \lesseqgtr 0$. Now

$$\frac{d^2y}{dx^2} = \frac{y}{[F(x)]^2}\left[(x - P)^2 + F(x) - (x - P)\frac{d}{dx}F(x)\right].$$

Whence if y is finite and $\neq 0$, $\frac{d^2y}{dx^2}$ does not change sign at $x = r$ because it is possible to select a neighborhood such that

$$|(x - P)^2| > \left|F(x) - (x - P)\frac{d}{dx}F(x)\right|$$

for an x differing from r by ϵ where ϵ is a small positive quantity. Therefore case (b) is not possible and $y = 0$ when a real root is repeated an even number of times. That is to say the range of the curve is limited at a point where a real root is repeated an even number of times. Thus Theorem I always holds for repeated roots.

For Theorem II it is clear that this Theorem holds for repeated roots when a non-repeated root lies closer to P, and on the same side, than the repeated root. Suppose that the repeated root is the nearest root to P (on a given side of P). Then by partial fractions:

$$\frac{dy}{y} = \frac{k_{11}\,dx}{(x - r_{11})} + \frac{k_{12}\,dx}{(x - r_{11})^2} + \frac{k_{13}\,dx}{(x - r_{11})^3} + \cdots + \frac{k_{41}\,dx}{(x - r_{41})} + \frac{k_{42}\,dx}{(x - r_{42})}$$
$$+ \cdots + \frac{k_{21}\,dx}{(x + r_{21})^2 + r_{0_1}^2} + \frac{k_{22}\,dx}{(x + r_{22})^2 + r_{0_2}^2} + \cdots + \frac{2k_{31}(x + r_{21})\,dx}{(x + r_{21})^2 + r_{0_1}^2} + \cdots$$

and on integrating:

$$y = C(x - r_{11})^{k_{11}}(x - r_{41})^{k_{41}}(x - r_{42})^{k_{42}} \cdots \left[(x + r_{21})^2 + r_{0_1}^2\right]^{k_{31}}$$
$$\cdots e^{k_{21}\arctan\frac{x + r_{21}}{r_{0_1}} + \cdots - \frac{k_{12}}{(x - r_{11})} - \frac{k_{13}}{2(x - r_{11})^2} - \cdots}.$$

Hence y can $= 0$ only for $x = r_{11}$ or for $x = r_{41}, r_{42}, \cdots$ and for no other finite values of x. Since by hypothesis y is bell-shaped, then the proper k_{ij} must be positive and Theorem II always holds for repeated roots.

Theorems I and II can now be combined and generalized in the form:

THEOREM: If $F(x)$ is a polynomial with real coefficients and $y = G(x)$ is a bell-shaped curve which is a solution of the differential equation

$$\frac{dy}{dx} = \frac{y(x - P)}{F(x)},$$

then the necessary and sufficient condition: that y be of unlimited range in both directions is that $F(x)$ have no real roots; that y be of limited range in one direction is that all of the real roots of $F(x)$ lie on the same side of P; that y be of limited range in both directions is that at least one real root of $F(x)$ lie on one side of P and one on the other.

COROLLARY: $F(x)$ must be negative throughout the range of y.

Suppose now that we have some statistics which we wish to graduate and the statistics are of such nature that we would expect a bell-shaped curve, rather than a J- or U-shaped curve, and we desire the best fit: If we use a curve which is a solution of the differential equation

$$\frac{dy}{dx} = \frac{y(x - P)}{F(x)}$$

(the Pearson Curves being special cases) to fit the statistics and if in computing the constants for the curve one of the following cases arise:

(a) $b_0' \not< 0$ when this constant is computed,

or (b) $B_0 \not< 0$ when the origin is moved to the mode,

or (c) a root is located within the range of the statistics then it means that:

1. A mistake may have been made in the computation: thus the Theorem just established provides a rough check on the work of computation,

2. If no mistake has been made in the computation it may indicate that the bell-shaped Pearson Curves will not closely fit the statistics and that some other graduation curves be used, e.g. the Gram-Charlier Types A or B might be tried,

3. If no mistake has been made in the computation it may happen that one of the bell-shaped Pearson Curves will give an excellent fit but a different method than or a modification of the Method or Moments should be used in order to compute the constants.

Part 3. Computing the Constants

At present, the constants of a frequency curve are computed as follows: First the moments are computed about an arbitrary origin, then the moments about the A.M. are determined, then β_1 and β_2 and the criterion are computed, after which the type of curve can be selected. From this point a separate procedure is followed for each curve. Now in the above method one will not know whether a root has been located in the range of statistics or not.

Take Pearson's differential equation

$$\frac{dy}{dx} = \frac{y(x - P)}{b_2x^2 + b_1x + b_0}.$$

Put $X = x - P$. Then $dX = dx$ and $x = X + P$, and

$$\frac{dy}{dx} = \frac{yX}{b_2(X + P)^2 + b_1(X + P) + b_0} = \frac{yX}{b_2X^2 + 2Pb_2X + b_1X + P^2b_2 + Pb_1 + b_0}.$$

Now put

$$b_2 = B_2$$
$$2Pb_2 + b_1 = B_1$$
$$P^2b_2 + Pb_1 + b_0 = B_0 .$$

Then we have

$$\frac{dy}{dX} = \frac{yX}{B_2X^2 + B_1X + B_0} \quad \text{or} \quad \frac{dy}{dx} = \frac{y(x - P)}{B_2(x - P)^2 + B_1(x - P) + B_0} . \tag{1}$$

It should be noted that for a particular curve, B_2, B_1 and B_0 are constants; i.e., their values do not change with a change of the origin. The values of b_1 and b_0 do change with a change in the origin.

If we clear equation (1) of fractions, multiply by $e^{\eta x}$ and integrate with respect to x over the range from x_1 to x_2, where

$$e^{\lambda_1\eta + \frac{\lambda_2\eta^2}{2!} + \frac{\lambda_3\eta^3}{3!} + \cdots} \equiv \int_{x_1}^{x_2} e^{\eta x}\, y dx ,$$

then successively differentiate with respect to η, and equate coefficients of like powers of η, we finally obtain:

$$\left.\begin{aligned}
&\lambda_1 - P + B_1 - 2PB_2 + 2B_2\lambda_1 = 0 ,\\
&\lambda_2 + B_0 - PB_1 + P^2B_2 + B_1\lambda_1 - 2PB_2\lambda_1 + 3B_2\lambda_2 + B_2\lambda_1^2 = 0 ,\\
&\lambda_3 + 2\lambda_2B_1 - 4PB_2\lambda_2 + 4B_2\lambda_3 + 4B_2\lambda_1\lambda_2 = 0 ,\\
&\lambda_4 + 3B_1\lambda_3 - 6PB_2\lambda_3 + 5B_2\lambda_4 + 6B_2\lambda_2^2 + 6B_2\lambda_1\lambda_3 = 0 .
\end{aligned}\right\} \tag{2}$$

Since we can compute the moments from the raw statistics and the semi-invariants from the moments, we may regard λ_2, λ_3 and λ_4 in these equations as knowns and the B_0, B_1, B_2, P and λ_1 as unknowns. But the origin has not yet been specified. Let the origin be placed at the A.M. where $\mu_1 = \lambda_1 = 0$. As λ_2, λ_3, λ_4, B_0, B_1 and B_2 are unchanged by a change of origin, we have:

$$\left.\begin{aligned}
&B_1 - P_0 - 2P_0B_2 = 0 .\\
&\lambda_2 + B_0 - P_0B_1 + P_0^2B_2 + 3B_2\lambda_2 = 0 ,\\
&\lambda_3 + 2B_1\lambda_2 - 4P_0B_2\lambda_2 + 4B_2\lambda_3 = 0 ,\\
&\lambda_4 + 3B_1\lambda_3 - 6P_0B_2\lambda_3 + 5B_2\lambda_4 + 6B_2\lambda_2^2 = 0 .
\end{aligned}\right\} \tag{3}$$

Now put

$$\left.\begin{aligned}
b_0' &= B_0 - P_0B_1 + P_0^2B_2 ,\\
b_1' &= B_1 - 2P_0B_2 ,\\
b_2' &= B_2 ;
\end{aligned}\right\} \tag{4}$$

then

$$\left.\begin{aligned} &b_1' - P_0 = 0, \\ &\lambda_2 + b_0' + 3b_2'\lambda_2 = 0, \\ &\lambda_3 + 2b_1'\lambda_2 + 4b_2'\lambda_3 = 0, \\ &\lambda_4 + 3b_1'\lambda_3 + 5b_2'\lambda_4 + 6b_2'\lambda_2^2 = 0. \end{aligned}\right\} \tag{5}$$

By reversing the transformation (4) we get:

$$\left.\begin{aligned} B_2 &= b_2', \\ B_1 &= b_1' + 2P_0b_2' \\ B_0 &= b_0' + P_0(b_1' + P_0b_2'). \end{aligned}\right\} \tag{6}$$

Now the above theory suggests the following procedure for computing the constants of a frequency curve: First the moments are computed about an arbitrary origin, then the semi-invariants are computed (or alternatively the moments about the A.M., either step involves about the same amount of work), then the equations (5) are solved and then by means of equations (6) the B_2, B_1 and B_0 are computed. Next solve the quadratic equation

$$B_2X^2 + B_1X + B_0 = 0.$$

The character of the roots of this equation indicates which type to use and it is unnecessary to compute the criterion. The constants of the frequency curve are simple functions of the roots of the above quadratic equation and can be readily found by integrating the diff. eq. (1) being careful to write the solution as a function of $X = x - P$. The rough checks mentioned in Part 2 can be quickly and conveniently applied when this procedure is followed.

GEORGE WASHINGTON UNIVERSITY.

A RECONSIDERATION OF SHEPPARD'S CORRECTIONS

By W. T. Lewis[1]

In computing the moments of a frequency distribution it is customary to find first what are known as the raw moments. These are obtained on the assumption that all the material of each class interval is concentrated at the middle point of the interval. It introduces what is called a grouping error because in fact the material does not all lie at the middle point. To compensate for this error W. F. Sheppard[2] derived a set of corrections. The hypothesis underlying his method is that the distribution may be regarded as similar to one to which the Euler-MacLaurin summation formula without its end terms may be applied. He presupposed such a curve, found its true moments, and then the raw moments that would be obtained if its area were concentrated at several equidistant abscissae. The relationship between these raw moments and the true moments of the curve furnished him with the corrections required for that distribution. If now our observed distribution may be supposed to be sufficiently like that one, we may use his corrections also on the observed data. One may note four points of criticism.

(1) The given distribution may not be similar to the one suggested, in the sense that it would be close to such a curve if the intervals of grouping were made very small; or at all events the purpose of finding the moments may be in part to decide whether or not it would become such a curve, and so one would not like to assume that to be true at the outset. A special case of importance in which this last is true occurs when one is finding the moments of a sample in order to determine whether it may have been drawn from a presupposed universe. It is inexact to use raw moments but it is illogical to use corrections that have been proved only for the universe being tested.

(2) Sheppard's argument does not make use of the one certain fact that is given in the hypothesis, viz: that the partial area of the given distribution over each class interval is exactly as stated. In fact, if, following the argument of some authors, the given curve be assumed to be exponential, it obviously cannot have partial areas everywhere exactly equal to the several given frequencies, for in particular its partial area is not zero beyond the given range.

(3) It is common to find distributions which do not have high contact at the ends of the range and for them Sheppard's corrections certainly fail. To obviate this criticism new corrections have been derived by Pairman and Pear-

[1] With the assistance of Burton H. Camp.

[2] The true values are given on page 220 of "Mathematical Part of Elementary Statistics, by Camp, D. C. Heath and Company, 1931.

son for the so-called abrupt cases. These new corrections are adequate to care for the abrupt cases but involve so much computation that it is a fair question whether it would not be simpler, first to distribute the given material over each interval by a smoothing process, and then to find without corrections the moments of the smoothed distribution.

(4) Even if one admits Sheppard's method in general, waiving the dubious question as to whether it is proper to start with an assumed curve instead of starting with the given distribution, it is doubtful whether there are any curves which have exactly the properties required. The high contact hypothesis may be put in different language as follows: using the notation of the Handbook[3] page 92, let $f(x)$ be the curve and x_t be the middle point of the slice. It is assumed that

$$\sum_t c x_t^r f^{(i)}(x_t) = \int_{-\infty}^{\infty} x^r f^{(i)}(x)\, dx\,; \qquad i = 0, 1, \cdots; \qquad r = 0, 1, \cdots;$$

c being the class interval. This means that if the moments of the curve be found by using *mid-ordinates times class interval*, instead of *areas*, one will obtain exactly the true moments of the curve, and that this will remain true for all the curves which are derivatives of this curve. This property is certainly not true of the normal curve; but it is almost true when r and the class interval are both small, and it is probably due to this fact that Sheppard's corrections seem to be good in practice.

Moreover, this high contact hypothesis cannot be true for any function over a limited range if the function is developable in Taylor's series about one end of the range. For the only function which has the required properties is identically zero, since the function and all its derivatives are required to vanish at that end of the range.

The primary purpose of this paper, therefore is to derive corrections similar to Sheppard's with a different set of assumptions. The results may be used as an approximate substitute for both Sheppard's and Pairman's. That is, they will apply approximately to both extreme cases and to the intermediate cases; on the whole they give better results than Sheppard's and are not so difficult to administer as Pairman's.

The argument runs as follows. When a distribution is given merely by class intervals, there is no way of knowing exactly what the distribution would have been had the class intervals been smaller; we do not know that we have a sample from an exponential curve, and even if we did we would not know that this sample would lie close to the exponential in form. We shall, however, try to draw a graduating curve in such a manner that (a) its partial area over each class interval will equal the frequency of the given distribution over that interval; and (b) its form within each class interval will be such that it will pass smoothly into the adjacent portions to the right and left. A good way to do this is by a

[3] H. L. Rietz, "Handbook of Math. Stat." Houghton Mifflin Co. (1924).

freehand graph, frankly recognizing that there are many forms that will do equally well. To obtain a numerical result it is necessary to use the equation of some curve. Again frankly recognizing that there are many types which will do equally well we choose the simplest to handle:

$$y = a + bt + ct^2.$$

Let the relative frequency distribution be defined by $f(i)$, $-m \leqq i \leqq n$, m, n, i being integers. To satisfy (a) we have the equation

$$\int_{i-\frac{1}{2}}^{i+\frac{1}{2}} y\,dt = f(i).$$

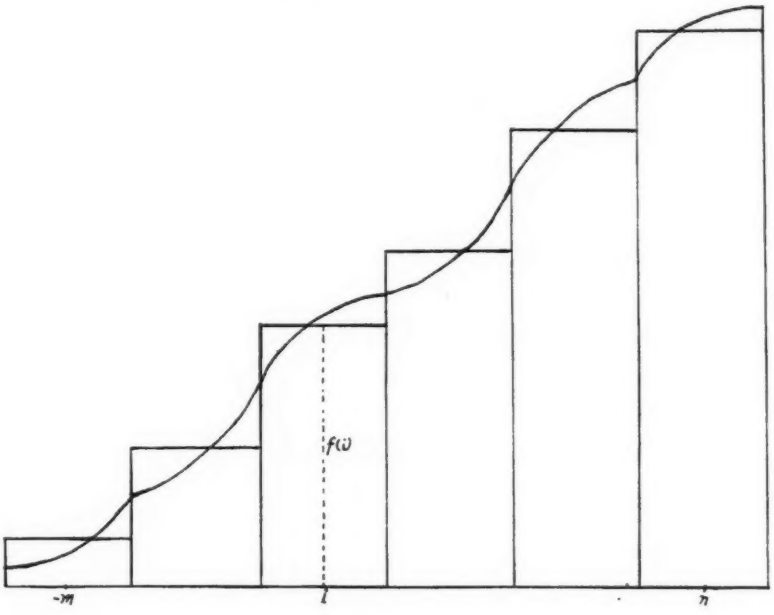

FIG. 1

To satisfy (b) we shall let

$$y = \tfrac{1}{2}[f(i) + f(i+1)] \text{ if } t = i + \tfrac{1}{2}.$$

The latter will hold for all values of i from $-m$ to $n-1$ inclusive, but the end intervals require special treatment. Here in order to satisfy as well as possible both the high contact and the abrupt cases, we wish to let the material be distributed according to the way the curve is behaving over the two nearest intervals on the right (at n) or left (at $-m$) rather than by the addition of zero frequencies beyond the given limits. To do this we let the slope of the parabolas be zero at the extremes:

$$\frac{dy}{dt} = 0 \qquad \text{at } t = -m - \tfrac{1}{2} \text{ and } t = n + \tfrac{1}{2}.$$

Then, if for example the frequencies are increasing as one nears the right end interval, the curve will rise over the right end interval; if they are decreasing, it will fall. These three conditions are sufficient to determine a continuous curve of the sort indicated in the figure. The exact moments of the curve may be found by integration and expressed in terms of the raw moments. The details are tedious and of an elementary nature and will be given only for the mean value $\bar{\nu}_1$.

To determine the coefficients of the parabola $y = a + bt + ct^2$ for the rectangle at $t = i$ we may write the following three equations; the first complying with the requirement that the area under the parabola from $t = i - \frac{1}{2}$ to $t = i + \frac{1}{2}$ equals the area of the rectangle at $t = i$, the second and third giving the ordinates at $i - \frac{1}{2}$ and $i + \frac{1}{2}$ respectively:

$$f(i) = \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} (a + bt + ct^2)dt\,,$$

$$\frac{f(i) + f(i+1)}{2} = a + b\,(i + \tfrac{1}{2}) + c\,(i + \tfrac{1}{2})^2\,,$$

$$\frac{f(i) + f(i-1)}{2} = a + b\,(i - \tfrac{1}{2}) + c\,(i - \tfrac{1}{2})^2\,.$$

Solving these three simultaneous equations we get for a, b, and c:

$$a = (\tfrac{5}{4} - 3i^2)\,f(i) + \left(\frac{3i^2}{2} - \frac{i}{2} - \frac{1}{8}\right) f(i+1) + \left(\frac{3i^2}{2} + \frac{i}{2} - \frac{1}{8}\right) f(i-1)\,,$$

$$b = 6if(i) + (\tfrac{1}{2} - 3i)f(i+1) - (\tfrac{1}{2} + 3i)f(i-1)\,,$$

$$c = -3f(i) + \tfrac{3}{2}f(i+1) + \tfrac{3}{2}f(i-1)\,,$$

and these hold for $-m + 1 \leqq i \leqq n - 1$.

For the parabola $y = a_1 + b_1 t + c_1 t^2$ over the first rectangle, i.e., where $i = -m$, we get the equations:

$$f(-m) = \int_{-m-\frac{1}{2}}^{-m+\frac{1}{2}} (a_1 + b_1\,t + c_1\,t^2)dt\,,$$

$$\frac{f(-m) + f(-m+1)}{2} = a_1 + b_1\,(-m + \tfrac{1}{2}) + c_1\,(-m + \tfrac{1}{2})^2\,,$$

$$b_1 + 2\,c_1\,(-m - \tfrac{1}{2}) = 0\,,$$

and their solutions:

$$a_1 = \tfrac{3}{4}\,(m^2 + m - \tfrac{1}{12})\,f(-m+1) - \tfrac{3}{4}\,(m^2 + m - \tfrac{17}{12})\,f(-m)\,,$$

$$b_1 = \tfrac{3}{4}\,(2m+1)\,f(-m+1) - \tfrac{3}{4}\,(2m+1)\,f(-m)\,,$$

$$c_1 = \tfrac{3}{4}\,f(-m+1) - \tfrac{3}{4}\,f(-m)\,.$$

Similarly for the parabola $y = a_n + b_n t + c_n t^2$ through the last rectangle at $i = n$ we get

$$f(n) = \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} (a_n + b_n t + c_n t^2) dt ,$$

$$\frac{f(n) + f(n-1)}{2} = a_n + b_n (n - \tfrac{1}{2}) + c_n (n - \tfrac{1}{2})^2 ,$$

$$b_n + 2 c_n n + c_n = 0 ,$$

and for the constants

$$a_n = \tfrac{3}{4} (n^2 + n - \tfrac{1}{12}) f(n-1) - \tfrac{3}{4} (n^2 + n - \tfrac{17}{12}) f(n) ,$$

$$b_n = - \tfrac{3}{4} (1 + 2n) f(n-1) + \tfrac{3}{4} (1 + 2n) f(n) ,$$

$$c_n = \tfrac{3}{4} f(n-1) - \tfrac{3}{4} f(n) .$$

Having obtained the constants for the graduating curve we will determine the moments of this curve in terms of those of the given frequency distribution.

Notation: Let the class interval be $c = 1$; let $\nu_s = \sum_{i=-m}^{n} i^s f(i)$ be the uncorrected s^{th} moment of the given frequency distribution about the given origin; let $\mu_s = \sum_{i=-m}^{n} (i - \nu_1)^s f(i)$ be the uncorrected s^{th} moment of the given frequency distribution about its uncorrected mean; let $\bar{\nu}_s$ be the corrected value of the s^{th} moment about the given origin; and let $\bar{\mu}_s$ be the corrected value of the s^{th} moment about the corrected mean. Thus ν_s and μ_s apply to the rectangles, and $\bar{\nu}_s$ and $\bar{\mu}_s$ apply to the curves as follows:

$$\bar{\nu}_s = \sum_{i=-m+1}^{n-1} \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} t^s (a + bt + ct^2) dt + \int_{-m-\frac{1}{2}}^{-m+\frac{1}{2}} t^s (a_1 + b_1 t + c_1 t^2) dt$$

$$+ \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} t^s (a_n + b_n t + c_n t^2) dt ,$$

$$\bar{\mu}_s = \sum_{i=-m+1}^{n-1} \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} (t - \bar{\nu}_1)^s (a + bt + ct^2) dt + \int_{-m-\frac{1}{2}}^{-m+\frac{1}{2}} (t - \bar{\nu}_1)^s (a_1 + b_1 t + c_1 t^2) dt$$

$$+ \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} (t - \bar{\nu}_1)^s (a_n + b_n t + c_n t^2) dt .$$

Using these symbols we have for the first moment about the given origin:

$$\bar{\nu}_1 = \sum_{i=-m+1}^{n-1} \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} t(a + bt + ct^2) dt + \int_{-m-\frac{1}{2}}^{-m+\frac{1}{2}} t(a_1 + b_1 t + c_1 t^2) dt$$

$$+ \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} t(a_n + b_n t + c_n t^2) dt$$

$$= \sum_{-m+1}^{n-1} \left[ai + b\left(i^2 + \frac{1}{12}\right) + c\left(i^3 + \frac{i}{4}\right)\right]$$
$$+ \left[- a_1 m + b_1\left(m^2 + \frac{1}{12}\right) - c_1\left(m^3 + \frac{m}{4}\right)\right]$$
$$+ \left[a_n n + b_n\left(n^2 + \frac{1}{12}\right) + c_n\left(n^3 + \frac{n}{4}\right)\right].$$

Substituting the values for the constants this becomes

$$\bar{\nu}_1 = \sum_{-m+1}^{n-1} \left\{ i\left[\left(\frac{5}{4} - 3i^2\right) f(i) + \left(\frac{3i^2}{2} - \frac{i}{2} - \frac{1}{8}\right) f(i+1) \right.\right.$$
$$\left. + \left(\frac{3i^2}{2} + \frac{i}{2} - \frac{1}{8}\right) f(i-1)\right]$$
$$+ (i^2 + \tfrac{1}{12})\,[6if(i) + (\tfrac{1}{2} - 3i)\,f(i+1) - (\tfrac{1}{2} + 3i)\,f(i-1)]$$
$$\left. + \left(i^3 + \frac{i}{4}\right)[-3f(i) + \tfrac{3}{2} f(i+1) + \tfrac{3}{2} f(i-1)]\right\}$$
$$+ \{- m\,[\tfrac{3}{4}\,(m^2 + m - \tfrac{1}{12})\,f(-m+1) - \tfrac{3}{4}\,(m^2 + m - \tfrac{17}{12})\,f(-m)]$$
$$+ (m^2 + \tfrac{1}{12})\,[\tfrac{3}{4}\,(2m+1)\,f(-m+1) - \tfrac{3}{4}\,(2m+1)\,f(-m)]$$
$$- \left(m^3 + \frac{m}{4}\right)[\tfrac{3}{4} f(-m+1) - \tfrac{3}{4} f(-m)]\}$$
$$+ \{n[\tfrac{3}{4}\,(n^2 + n - \tfrac{1}{12})\,f(n-1) - \tfrac{3}{4}\,(n^2 + n - \tfrac{17}{12})\,f(n)]$$
$$+ \left(n^2 + \frac{1}{12}\right)\left[-\frac{3}{4}\,(1+2n)\,f(n-1) + \frac{3}{4}\,(1+2n)\,f(n)\right]$$
$$\left. + \left(n^3 + \frac{n}{4}\right)\left[\frac{3}{4} f(n-1) - \frac{3}{4} f(n)\right]\right\}.$$

$$\bar{\nu}_1 = \sum_{-m+1}^{n-1} \left[if(i) + \frac{1}{24} f(i+1) - \frac{1}{24} f(i-1)\right] + \frac{1}{16} f(-m+1)$$
$$- \left(m + \frac{1}{16}\right) f(-m) - \frac{1}{16} f(n-1) + \left(n + \frac{1}{16}\right) f(n).$$

$$\sum_{-m+1}^{n-1} if(i) = \sum_{-m}^{n} if(i) - (-m)\,f(-m) - nf(n) = \nu_1 + mf(-m) - nf(n).$$

$$\frac{1}{24} \sum_{i=-m+1}^{n-1} f(i+1) = \frac{1}{24} \sum_{k=-m+2}^{n} f(k) = \frac{1}{24} \sum_{-m}^{n} f(k) - \frac{1}{24} f(-m+1) - \frac{1}{24} f(-m)$$
$$= \frac{1}{24} - \frac{1}{24} f(-m+1) - \frac{1}{24} f(-m).$$

$$\frac{1}{24}\sum_{i=-m+1}^{n-1} f(i-1) = \frac{1}{24}\sum_{j=-m}^{n-2} f(j) = \frac{1}{24}\sum_{-m}^{n} f(j) - \frac{1}{24}f(n-1) - \frac{1}{24}f(n)$$

$$= \frac{1}{24} - \frac{1}{24}f(n-1) - \frac{1}{24}f(n).$$

$$\bar{\nu}_1 = \nu_1 + mf(-m) - nf(n) + \frac{1}{24} - \frac{1}{24}f(-m+1) - \frac{1}{24}f(-m) - \frac{1}{24}$$

$$+ \frac{1}{24}f(n-1) + \frac{1}{24}f(n) + \frac{1}{16}f(-m+1)$$

$$- \left(m + \frac{1}{16}\right)f(-m) - \frac{1}{16}f(n-1) + \left(n + \frac{1}{16}\right)f(n).$$

$$\bar{\nu}_1 = \nu_1 - \frac{5}{48}f(-m) + \frac{5}{48}f(n) + \frac{1}{48}f(-m+1) - \frac{1}{48}f(n-1).$$

Using this same notation and method for the higher moments we get

$$\bar{\mu}_2 = \nu_2 - \frac{1}{12} - \bar{\nu}_1^2 + \left(\frac{5m}{24} + \frac{7}{80}\right)f(-m) + \left(\frac{5n}{24} + \frac{7}{80}\right)f(n)$$

$$+ \left(\frac{-m}{24} - \frac{1}{240}\right)f(-m+1) + \left(\frac{-n}{24} - \frac{1}{240}\right)f(n-1).$$

$$\bar{\mu}_3 = \nu_3 - 3\bar{\nu}_1\bar{\mu}_2 - \frac{\bar{\nu}_1}{4} - \bar{\nu}_1^3 + f(-m)\left[\frac{-5}{16}m^2 - \frac{21}{80}m - \frac{17}{120}\right]$$

$$+ f(n)\left[\frac{5}{16}n^2 + \frac{21}{80}n + \frac{17}{120}\right] + f(-m+1)\left[\frac{m^2}{16} + \frac{m}{80} + \frac{1}{120}\right]$$

$$+ f(n-1)\left[\frac{-n^2}{16} - \frac{n}{80} - \frac{1}{120}\right].$$

$$\bar{\mu}_4 = \nu_4 - 4\bar{\mu}_3\bar{\nu}_1 - 6\bar{\mu}_2\bar{\nu}_1^2 - \bar{\nu}_1^4 - \frac{\bar{\mu}_2}{2} - \frac{\bar{\nu}_1^2}{2} - \frac{17}{64}$$

$$+ f(-m)\left[\frac{5m^3}{12} + \frac{21m^2}{40} + \frac{17m}{30} + \frac{313}{1680}\right] + f(n)\left[\frac{5n^3}{12} + \frac{21n^2}{40} + \frac{17n}{30} + \frac{313}{1680}\right]$$

$$+ f(-m+1)\left[\frac{-m^3}{12} - \frac{m^2}{40} - \frac{m}{30} - \frac{1}{336}\right] + f(n-1)\left[\frac{-n^3}{12} - \frac{n^2}{40} - \frac{n}{30} - \frac{1}{336}\right].$$

SPECIAL CASES

The above formulae are rather long and in practice the special cases below will frequently be preferred.

(*a*) We may usually take the origin at or very near the middle of the range so that $m = n$, at least approximately.

If $m = n$:

$$\bar{\nu}_1 = \nu_1 - \frac{5}{48}f(-m) + \frac{5}{48}f(n) + \frac{1}{48}f(-m+1) - \frac{1}{48}f(n-1).$$

$$\bar{\mu}_2 = \nu_2 - \frac{1}{12} - \bar{\nu}_1^2 + \left(\frac{5m}{24} + \frac{7}{80}\right)[f(-m) + f(n)] + \left(\frac{-m}{24} - \frac{1}{240}\right)[f(-m+1) + f(n-1)].$$

$$\bar{\mu}_3 = \nu_3 - 3\bar{\nu}_1\bar{\mu}_2 - \frac{\bar{\nu}_1}{4} - \bar{\nu}_1^3 + \left[\frac{5m^2}{16} + \frac{21m}{80} + \frac{17}{120}\right][f(n) - f(-m)] + \left[\frac{m^2}{16} + \frac{m}{80} + \frac{1}{120}\right][f(-m+1) - f(n-1)].$$

$$\bar{\mu}_4 = \nu_4 - 4\bar{\mu}_3\bar{\nu}_1 - 6\bar{\mu}_2\bar{\nu}_1^2 - \bar{\nu}_1^4 - \frac{\bar{\mu}_2}{2} - \frac{\bar{\nu}_1^2}{2} - \frac{17}{64} + \left[\frac{-m^3}{12} - \frac{m^2}{40} - \frac{m}{30} - \frac{1}{336}\right][f(-m+1) + f(n-1)] + \left[\frac{5m^3}{12} + \frac{21m^2}{40} + \frac{17m}{30} + \frac{313}{1680}\right][f(-m) + f(n)].$$

(*b*) Except in the abrupt cases the end frequencies and the difference between those next to the ends will be so small (relative to unity) that they will have a negligible effect on the corrections. If $m = n$ as in (*a*), and if also

$$f(-m) = f(n) = 0 \text{ and } f(-m+1) - f(n-1) = 0:$$

$$\bar{\nu}_1 = \nu_1.$$

$$\bar{\mu}_2 = \nu_2 - \bar{\nu}_1^2 - \frac{1}{12} + f(-m+1)\left[\frac{-m}{12} - \frac{1}{120}\right].$$

$$\bar{\mu}_3 = \nu_3 - 3\bar{\nu}_1\bar{\mu}_2 - \frac{\bar{\nu}_1}{4} - \bar{\nu}_1^3.$$

$$\bar{\mu}_4 = \nu_4 - 4\bar{\mu}_3\bar{\nu}_1 - 6\bar{\mu}_2\bar{\nu}_1^2 - \bar{\nu}_1^4 - \frac{\bar{\mu}_2}{2} - \frac{\bar{\nu}_1^2}{2} - \frac{17}{64} + f(-m+1)\left[\frac{-m^3}{6} - \frac{m^2}{20} - \frac{m}{15} - \frac{1}{168}\right].$$

These formulae have been written in the form which makes the computing simple. The following makes a comparison with Sheppard's corrections easy.

$$\bar{\nu}_1 = \nu_1 .$$

$$\bar{\mu}_2 = \mu_2 - \frac{1}{12} + f(-m+1)\left[\frac{-m}{12} - \frac{1}{120}\right].$$

$$\bar{\mu}_3 = \mu_3 + \nu_1\left(\frac{m}{4} + \frac{1}{40}\right)f(-m+1).$$

$$\bar{\mu}_4 = \mu_4 - \frac{\mu_2}{2} - \frac{43}{192} + f(-m+1)\left[\frac{-m^3}{6} - \frac{m^2}{20} - \frac{m}{40} - \frac{1}{560} - \frac{m\nu_1^2}{2} - \frac{\nu_1^2}{20}\right].$$

The following special case is also useful in comparing my formulae with Sheppard's.

(c) Let $f(-m) = \frac{1}{5} f(-m+1)$ and $f(n) = \frac{1}{5} f(n-1)$. This produces a graduating curve which is exactly tangent to the t-axis at the ends of the range and is everywhere continuous—though it does not have continuous derivatives at certain isolated points. It is, however, a curve which to the eye cannot be distinguished from the type assumed in the Euler-MacLaurin theorem, which lies at the base of Sheppard's formulae. My corrections become:

$$\bar{\nu}_1 = \nu_1 ,$$

$$\bar{\mu}_2 = \mu_2 - \frac{1}{12} + \frac{1}{15}[f(-m) + f(n)],$$

$$\bar{\mu}_3 = \mu_3 - \frac{\nu_1}{5}[f(-m) + f(n)] + \left[\frac{-m}{5} - \frac{1}{10}\right]f(-m) + \left[\frac{n}{5} + \frac{1}{10}\right]f(n),$$

$$\bar{\mu}_4 = \mu_4 - \frac{\mu_2}{2} - \frac{43}{192} + \frac{2}{5}\left[(\nu_1 + m)^2 + \nu_1 + m + \frac{29}{84}\right]f(-m) + \frac{2}{5}\left[(\nu_1 - n)^2 - \nu_1 + n + \frac{29}{84}\right]f(n).$$

Sheppard's are:

$$\bar{\nu}_1 = \nu_1 ,$$

$$\bar{\mu}_2 = \mu_2 - \frac{1}{12},$$

$$\bar{\mu}_3 = \mu_3 ,$$

$$\bar{\mu}_4 = \mu_4 - \frac{\mu_2}{2} + \frac{7}{240}.$$

Let us compare my results with Sheppard's in the very special case in which $f(-m) = f(n) = 1/7$, $f(0) = 5/7$, $m = n = 1$. The odd moments vanish. My corrections for μ_2 and μ_4 are

$$\bar{\mu}_2 = 0.2214, \qquad \bar{\mu}_4 = 0.1870.$$

Sheppard's are

$$\bar{\mu}_2 = 0.2024, \qquad \bar{\mu}_4 = 0.1720.$$

The numerical difference between the $\bar{\mu}_2$'s is 0.0190, and the numerical difference between the $\bar{\mu}_4$'s is 0.0150.

This example shows that Sheppard's corrections are not valid to the precision to which they are usually given if they are to be used for the purpose of correcting raw moments. The last term in the fourth moment correction, 7/240, might equally well be, for example, $-43/192$ as in my special case. This will become more evident to the reader if he will draw the curve indicated in this example. To the eye it will appear exactly like the kind specified in the Euler-MacLaurin theorem; for example, much like the normal curve. Now suppose one adopted for the moment the point of view (which I have criticized earlier) of starting with the curve used in this example, breaking it up into three partial areas and then finding the relation between the true and the raw moments. The partial areas found would be exactly those used in this example and this method would give us Sheppard's corrections, but they would not be exactly correct, for in this instance my formulae give exactly the relationship between the true and the raw moments. The difference is due to the fact that in this instance the assumptions permitting the use of the Euler-MacLaurin theorem in abbreviated form are not justified for this curve. But there is no way of telling at the outset, if one has given initially only the partial areas, whether precisely this curve or another which to the eye would appear very much like it is truly the curve which will graduate the same material when subjected to a finer classification.

THE POINT BINOMIAL AND PROBABILITY PAPER

By Frank H. Byron[1]

1. An approximation to the sum of a number of consecutive terms of the point binomial may be found graphically and quite expeditiously by means of so-called "probability paper." This paper is ruled so that the (x, y) graph of the equation of the integral of the normal curve

$$y = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x}{2}} \, dx \tag{1}$$

is a straight line. Let the successive terms of the point binomial be represented as follows:

$$(p + q)^n = u_0 + u_1 + \cdots + u_t + \cdots + u_n, \tag{2}$$

where $u_t = {}_nC_t p^{n-t} q^t$ and $p \geqq q$. Then the (x, y) graph of the equation,

$$y = \sum_{i=0}^{t} u_i, \qquad t + \tfrac{1}{2} = x, \tag{3}$$

i.e., of the sum of first $(t + 1)$ terms of this point binomial, is, in all but extreme cases, a set of points lying on a gently turning curve, so gently that its form may be represented closely by two straight lines, each passing through the median point as will be explained in the next section. As paper of this sort is readily obtainable, and as this method yields as great accuracy as is really useful in many problems, it is suggested that its use ought to be quite general.

2. **Sheppard's Corrections.** The formulae for the moments of the point binomial, mean $= qn$, $\sigma^2 = pqn$, are exact without any corrections such as are used for grouped material. This fact has led us all (apparently) to assume that in fitting the curve to the point binomial one would get a better fit by equating the moments of the curve to the uncorrected moments of the point binomial rather than to the corrected moments. The studies made in connection with the preparation of this paper show that when the purpose is to equate areas to sums of terms the corrected moments should be used. The theoretical basis for this conclusion is as follows:

To simplify the argument let us suppose that one were seeking that curve of Charlier type,

$$F(x) = c_0\phi_0(x) + c_1\phi_1(x) + \cdots c_4\phi_4(x), \tag{4}$$

[1] With the assistance of Burton H. Camp.

(where ϕ_0 is the normal curve and ϕ_1, ϕ_2, $\cdots$ its successive derivatives) whose integral would best fit the graph of (3). Since fitting is required only at the isolated points $x = \frac{1}{2}, 1\frac{1}{2}, 2\frac{1}{2}, \cdots$, it is clear that one might obtain this by the two following steps. First let $f(x)$ be any function whose integral meets exactly the requirement at these isolated points. What values this integral has at other points does not for the moment concern us. There are an infinite number of such $f(x)$ curves. Next let the c's of (4) be so chosen that $F(x)$ will fit $f(x)$ as nearly as possible. The ordinary derivation of the c's supposes that the fit between $f(x)$ and $F(x)$ is to be made by least squares, the residuals being weighted by the factor $1/\sqrt{\phi(x)}$. No matter what $f(x)$ is chosen, the c's can be determined so that the weighted integral of $(f(x) - F(x))^2$ will be a minimum, but the value of this minimum will vary from one $f(x)$ to another. We now desire to select that $f(x)$ which will make this minimum value as small as possible, and it is reasonable to suppose that our best selection will be some $f(x)$ which is as kindred to the nature of $F(x)$ as possible. We shall not therefore choose an $f(x)$ which oscillates wildly between the points where perfect fitting is required, (Fig. 1) nor yet an $f(x)$ which is made up of the top bases of the point binomial

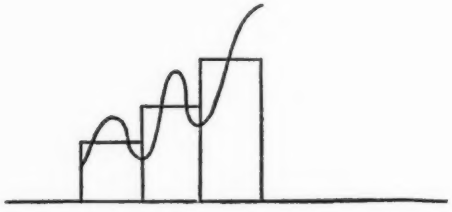

FIG. 1

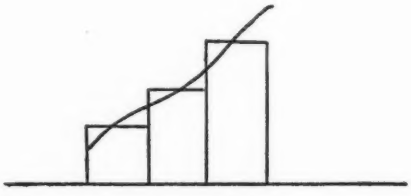

FIG. 2

histogram; we shall prefer a modification (Fig. 2) of that histogram by a smoothing process. Such an $f(x)$ will not have the exact moments of the point binomial, but, more nearly, those moments corrected for grouping. Then the determination of the c's will come out in terms of these corrected moments, not in terms of the uncorrected moments. (In fact the uncorrected moments would be the exact moments of an $f(x)$ having an oscillatory character between the important points.)

Of course, when n is large, the difference is too small to be noticed and the use of Sheppard's corrections is not worth while, and since n usually is large when approximations of this sort are needed, the point is not usually important. It was important in the computation of the tables of §4. Moreover, the use of Sheppard's corrections does not invariably yield better results, the gain being masked sometimes by other effects to be considered in §3. An excellent illustration of uniformly better results is in fitting $(\frac{1}{2} + \frac{1}{2})^9$ by a curve of Type 4. The errors in the sums as derived from (4) with and without the corrections, is given on the following page.

t	0	1	2	3	4	5	6	7	8	9
With Corrections	.0002	.0001	−.0003	−.0001	.0000	.0001	.0003	−.0001	−.0002	.0000
Without Corrections	.0007	.0022	.0039	.0036	.0000	−.0036	−.0039	−.0022	−.0007	−.0001

3. **The Stubby End.** The other effects which mask this improvement are especially noticeable at the stubby end of a point binomial. We have to keep in mind here that the approximating curve (such as (4)), is required to turn a sharp corner, for, due to the least square method of fitting, it is just as important that it be close to zero when t is negative, as it is that it be close to $u_0, u_1, \cdots$ when t is positive. Therefore, in order to turn this corner it has to dip below the x-axis in the neighborhood of $t = -\frac{1}{2}$. This makes the approximating curve too low just to the right of $t = -\frac{1}{2}$, unless the whole curve be arbitrarily widened. This arbitrary widening is customarily performed by not using Sheppard's correction for σ, and the result is a betterment of the fit at these points but a corresponding loss over the rest of the infinite interval. A good example[2] is $(\frac{2}{3} + \frac{1}{3})^{25}$. The fit is worse at the left end when Sheppard's corrections are used but better over the rest of the interval.

The same difficulty arises in another connection. If we compare the closeness of fit to a point binomial made by $F(x)$ as written in (4) and by $F(x)$ as it would be written if c_4 were zero, it often happens (as is well known) that the latter is actually slightly better on the average. How can this be true if the c's are chosen by the method of least squares and the best choice as thus indicated makes c_4 different from zero? The answer is that the c's are chosen so that the fit is best over the infinite interval, not merely over the interval from $t = -\frac{1}{2}$ to $t = n + \frac{1}{2}$, and that furthermore the distant points are weighted more heavily than those near the center. Thus it might happen that a choice, other than the least square choice, and one in which c_4 would be zero, might be better for the restricted interval covered by the point binomial. This does happen especially when due to the abruptness of the stubby end of a very skew binomial, the curve has to dip below the axis in order to get by a sharp corner. A good example is the problem considered by Fry:[3] $(\frac{9}{10} + \frac{1}{10})^{100}$. All the effects mentioned are present here. The fit is on the average a little worse if c_4 is not equal to zero over the point binomial interval, a little better over the infinite interval.

4. For graphical purposes a sufficiently good approximation to the median of $(p + q)^n$, is given by

$$M = nq - (p - q)/6.$$

[2] The true values are given on page 220 of Mathematical Part of Elementary Statistics, by Camp, D. C. Heath and Company, 1931.

[3] T. C. Fry, Probability and its Engineering Uses, p. 258, Van Nostrand, 1928.

The following tables enable us to find the first quartile Q_1, and the ninth decile D_9. The accuracy to which they can be plotted is only about one-tenth that to which they are given here. Therefore accurate interpolation is seldom necessary. The values of S_{t+1} are to be read from the graph at the points $t + \frac{1}{2}$, as indicated in the directions preceding the tables. The graphical method will be found efficient if one uses common sense in the computation. Numbers which are to be plotted should not be computed to a higher degree of accuracy than can be used graphically. In reading the values of S_{t+1} it is well to remember that the true values lie on a curve, and that outside the interval from Q_1 to D_9, they are slightly less than those given by the straight line. Once the graph has

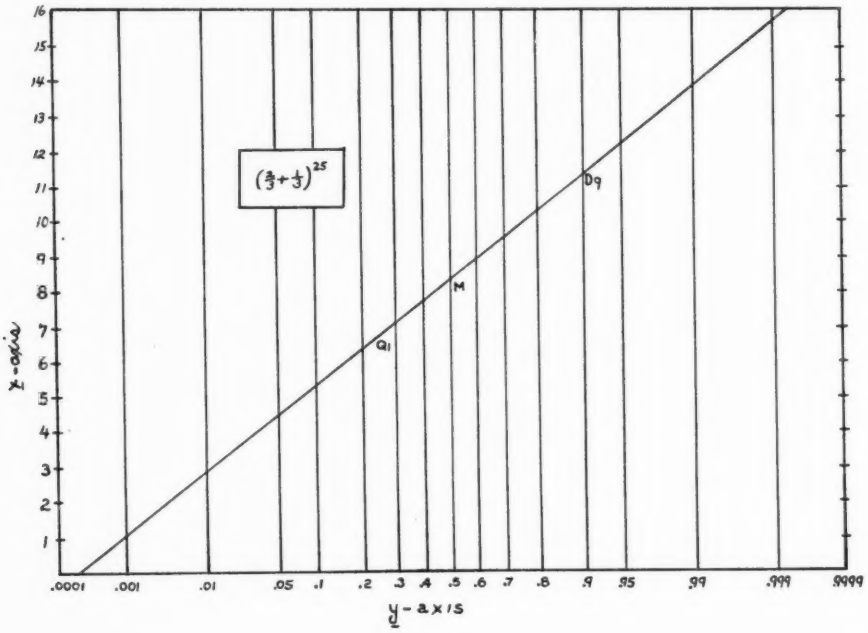

FIG. 3

been made, all the values of S_{t+1} can be read quickly; it is not necessary to make a separate computation for each t. This method is therefore specially advantageous when one wishes to find several sums of this sort for the same point binomial. It should also be noticed that one can tell from the appearance of the graph about how far the true sum would be from the two straight lines and so estimate the error to which his reading is liable.

5. **Illustration.** Find the sum of the first 7 terms of $(\frac{2}{3} + \frac{1}{3})^{25}$.

Here $t = 6$, $M = 8.278$, $Q_1 = 6.726$, $D_9 = 11.369$. The graph shows that $\sum_0^t = 0.224$. The true value is 0.222. So the error is 0.002.

An idea of the accuracy of the method is given by the errors (out of two places) that would be obtained for this point binomial for various values of t, as follows:

t	2	4	6	8	10	12	14	16
Errors	.00	.01	.00	.00	.00	.00	.00	.00

DIRECTIONS FOR USE OF THE TABLES: Let $p = q$, $M = nq - (p - q)/6$, $Q_1 = x_1 + qn$, $D_9 = x_2 + qn$. On the graph draw the lines MQ_1 and MD_9. Read S_{t+1} at $t + \frac{1}{2}$.

Values of x_1

p \ n	2000	1000	750	500	400	300	200	100	75	50	25
.99	−.693	−.701	−.704	−.710	−.714	−.720	−.728	−.747	−.756	−.771	−.804
.98	−.688	−.693	−.696	−.700	−.703	−.707	−.714	−.728	−.735	−.746	−.770
.97	−.685	−.690	−.692	−.696	−.698	−.701	−.707	−.718	−.724	−.734	−.784
.96	−.684	−.687	−.689	−.693	−.695	−.697	−.702	−.712	−.718	−.726	−.744
.95	−.683	−.686	−.688	−.691	−.692	−.695	−.699	−.708	−.713	−.721	−.737
.94	−.682	−.685	−.686	−.689	−.691	−.693	−.697	−.705	−.709	−.717	−.732
.93	−.681	−.684	−.685	−.688	−.689	−.691	−.695	−.703	−.707	−.713	−.727
.92	−.681	−.683	−.685	−.687	−.688	−.690	−.693	−.701	−.704	−.710	−.723
.91	−.680	−.683	−.684	−.686	−.687	−.689	−.692	−.699	−.702	−.708	−.720
.90	−.680	−.682	−.683	−.685	−.686	−.688	−.690	−.697	−.700	−.704	−.717
.88	−.679	−.681	−.682	−.684	−.685	−.686	−.689	−.695	−.697	−.702	−.713
.85	−.679	−.680	−.681	−.682	−.683	−.685	−.687	−.691	−.694	−.698	−.707
.80	−.677	−.679	−.679	−.681	−.681	−.682	−.684	−.688	−.690	−.693	−.700
.75	−.677	−.678	−.678	−.679	−.680	−.681	−.682	−.685	−.686	−.689	−.694
.70	−.676	−.677	−.677	−.678	−.679	−.679	−.680	−.682	−.683	−.685	−.690
.65	−.676	−.676	−.677	−.677	−.677	−.678	−.678	−.680	−.681	−.682	−.686
.60	−.675	−.676	−.676	−.676	−.676	−.677	−.677	−.678	−.679	−.680	−.682
.50	−.675	−.675	−.675	−.675	−.675	−.675	−.675	−.675	−.675	−.675	−.675

Values of x_2

p \ n	2000	1000	750	500	400	300	200	100	75	50	25
.99	1.307	1.318	1.325	1.336	1.344	1.356	1.378	1.439	1.481	See Auxiliary	
.98	299	307	311	318	323	330	343	376	396	Tables	
.97	295	301	304	310	314	319	329	353	367		
.96	293	298	301	306	309	313	321	341	352		
.95	292	296	299	303	305	309	316	332	342		
.94	291	295	297	300	303	306	312	327	335		
.93	290	293	295	298	301	304	309	322	329	1.342	1.374
.92	289	292	294	297	299	302	307	318	325	336	365
.91	289	292	293	296	298	300	305	315	321	331	357
.90	288	291	292	295	296	299	303	313	318	325	351
.88	287	290	291	293	295	297	300	309	313	321	341
.85	286	288	289	291	292	294	297	304	308	314	330
.80	285	287	288	289	290	291	293	298	301	306	317
.75	284	285	286	287	288	289	291	294	297	300	308
.70	284	285	285	286	286	287	288	291	293	295	301
.65	283	284	284	285	285	286	286	288	290	292	296
.60	283	283	283	284	284	284	285	286	287	288	291
.50	282	282	282	282	282	282	282	282	282	282	282

Auxiliary Table

p \ n	60	50	40	35	30	25	20
.99	1.525	1.575	1.663	1.740	1.871	2.149	3.209
.98	416	435	455	488	520	1.568	1.652
.97	381	394	413	433	445	472	514
.96	362	372	387	397	410	428	457
.95	350	359	370	378	389	405	425
.94	336	349	359	366	375	387	405

INEQUALITIES AMONG AVERAGES

By Nilan Norris

Numerous inequalities among averages of various types are condensed in the monotonic character of the function

$$\phi(t) = \left(\frac{x_1^t + x_2^t + \cdots + x_n^t}{n}\right)^{\frac{1}{t}}$$

of the positive numbers $x_1, x_2, \cdots, x_n$, not all equal each to each. For $t = -1$ this function is the harmonic mean; for $t = 0$ it is the geometric mean; for $t = 1$ the arithmetic mean; and for $t = 2$ the root mean square. The relations among these four means which customarily are proved by special and disconnected methods appear easily as applications of the theorem that $\phi(t)$ is an increasing function of t. That is, for any values of t_1 and t_2 such that $-\infty < t_1 < t_2 < +\infty$, it will be true that $\phi(t_1) < \phi(t_2)$. Several proofs of this theorem have been published, many of them very complex. An extremely simple proof is herewith presented.[1]

That $\phi(t)$, $\phi'(t)$ and $\phi''(t)$ all exist and are continuous for all real values of t may be shown by expanding each of the quantities x_i^t in a series of powers of t and considering the remainders after each of the first three terms. The ordinary rule for evaluating forms reducing to 0/0, which requires the function under consideration to be continuous and to have at least a continuous first derivative for $t = 0$, may then be applied to $[\log \phi(t)]/t$ to show that $\phi(0)$ is the geometric mean. It is clear that $\phi(-\infty)$ and $\phi(+\infty)$ are respectively the least and the greatest of the x_i. This fact and the monotonic property of $\phi(t)$ make it evident that for each real value of t, the function may be regarded as an average in the usual sense that it lies within the range of the observations.

For a simple demonstration of the increasing character of $\phi(t)$, consider the auxiliary function

$$F(t) = t^2 \frac{\phi'(t)}{\phi(t)} = t^2 \frac{d}{dt}\left\{\frac{1}{t} \log \frac{\Sigma x^t}{n}\right\} = t \frac{\Sigma x^t \log x}{\Sigma x^t} - \log \frac{\Sigma x^t}{n}.$$

It is clear that $\phi'(t)$ has the same sign as $F(t)$. The theorem will be proved by showing that the sign of $F(t)$ is positive for all values of t except zero, when $\phi'(t)$ vanishes.

---

[1] Professor Harold Hotelling rendered invaluable assistance in condensing for publication the material herein presented from a more extended study of generalized mean value functions.

Differentiating the last expression with respect to t, one obtains upon simplification

$$F'(t) = \frac{t}{(\Sigma x^t)^2}\left[(\Sigma x^t)(\Sigma x^t \log^2 x) - (\Sigma x^t \log x)^2\right].$$

By Cauchy's inequality (known as Schwarz' inequality when applied to integrals instead of sums), the expression in square brackets is positive. Hence $F'(t)$ has the same sign as t. Consequently $F(t)$, since it diminishes for negative values of t and increases for positive values, has a minimum for $t = 0$. But by direct substitution, $F(0) = 0$. It follows that $F(t)$ and $\phi'(t)$ are positive for all values of t other than zero. Therefore $\phi(t)$ is an increasing function.

By direct general methods it is possible to show that

$$\phi'(0) = (\Pi x)^{\frac{1}{n}} \frac{1}{2n^2}\left[n\Sigma(\log x)^2 - (\Sigma \log x)^2\right].$$

This expression obviously vanishes only when $n\Sigma(\log x)^2 = (\Sigma \log x)^2$, a condition which is satisfied only in the trivial case when $x_1 = x_2 = \cdots = x_n$.

A proof exactly parallel to that given above may be applied to integrals or, more generally, to Stieltjes integrals. The monotonic increasing character of $\left[\int_{x=0}^{\infty} x^t d\psi(x)\right]^{\frac{1}{t}}$ appears in this way if one assumes that $\psi(x)$ is a non-decreasing function integrable in the Riemann-Stieltjes sense, such that $\psi(\infty) - \psi(0) = 1$, and such that $\int_{x=0}^{\infty} x^t d\psi(x)$ exists for every real value of t. In terms of statistical theory, this consideration extends the theorem from samples to populations of a very general character.

Proof of the increasing character of $\phi(t)$ has also been derived from Hölder's inequality, the demonstration being expressed in terms of Stieltjes integrals.[2] The simplest general proof of the monotonic attribute of $\phi(t)$ heretofore published appears to be that of Paul Lévy.[3] As early as 1840 Bienaymé[4] presented a generalized form of $\phi(t)$, namely,

$$\left(\frac{c_1 a_1^m + c_2 a_2^m + \cdots + c_n a_n^m}{c_1 + c_2 + \cdots + c_n}\right)^{\frac{1}{m}},$$

and announced, without proof, its increasing character. In 1858 a proof of the monotonic quality of $\phi(t)$ for special cases was published by Schlömilch.[5] Of

---

[2] J. Shohat, "Stieltjes Integrals in Mathematical Statistics," *Annals of Mathematical Statistics* (American Statistical Association, Ann Arbor, 1930), Vol. 1, No. 1, p. 84.

[3] *Calcul des Probabilités* (Gauthier-Villars et Cie., Paris, 1925), pp. 157 *f.*

[4] Jules Bienaymé, *Société Philomatique de Paris*, Extraits des Procès-Verbaux des Seances Pedant L'Anée 1840 (Imprimerie D'A. René et Cie., Paris, 1841), Seance du 13 juin 1840, p. 68.

[5] O. Schlömilch, "Ueber Mittelgrössen verschiedener Ordnungen," *Zeitschrift für Mathematik und Physik* (B. G. Teubner, Leipzig, 1858), Vol. 3, pp. 303 *f.*

the more recent general proofs of the increasing character of $\phi(t)$ which have appeared, those of Jensen,[6] Pólya,[7] Jessen,[8] and Carathéodory[9] may be mentioned. A recent application of $\phi(t)$ to index number theory is that of Professor John B. Canning.[10]

VASSAR COLLEGE.

---

[6] J. L. W. V. Jensen, "Sur Les Fonctions Convexes Et Les Inegalités Entre Les Valeurs Moyennes," *Acta Mathematica* (Beijers Bokförlagsaktielbolag, Stockholm, 1905), Vol. 30, pp. 183–185.

[7] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze Aus Der Analysis* (Julius Springer, Berlin, 1925), Vol. I, pp. 54 *f.* and 210.

[8] Børge Jessen, "Bemaerkninger om koveskse Funktioner og Uligheder imellem Middelvaerdier," *Matematisk Tidsskrift* (Charles Johansens Bogtrykkeri, Copenhagen, 1931), No. 2, 1931, pp. 26–28.

[9] Attributed to Professor Constantin Carathéodory in an unpublished manuscript of Professor Harold Hotelling.

[10] "A Theorem Concerning a Certain Family of Averages of a Certain Type of Frequency Distribution," a paper presented before a joint meeting of the American Statistical Association and the Econometric Society at Berkeley, California, June 22, 1934.

MATHEMATICAL EXPECTATION OF PRODUCT MOMENTS OF SAMPLES DRAWN FROM A SET OF INFINITE POPULATIONS

By Hyman M. Feldman[1]

Introduction

In the second part of his investigations, "On the Mathematical Expectation of Moments of Frequency Distributions,"[2] Tchouproff presented a method which may be interpreted as sampling from a set of infinite univariate populations. In the present paper this method is extended to the study of moments of product moments of samples drawn from a set of infinite bivariate populations. It is also shown how this method may be extended to populations of higher order by deriving some of the simpler formulae for populations of three and four variables.

Tchouproff's method has been criticised[3] because of the complicated algebra. On close examination it is found, however, that it is not the algebra which is complicated but rather the symbolism. Tchouproff introduced a great variety of symbols both in his derivations and in his results. As a consequence his work seems very intricate. If, however, the number of symbols is reduced, and the symbols themselves are simplified, which can be easily accomplished, the underlying idea of Tchouproff's method is found to be very simple.

Quite a complete study of product moments of any bivariate population has been made by Joseph Pepper in his "Studies in the Theory of Sampling."[4] His method is essentially an extension of Church's[5] method, in his studies of univariate populations, to bivariate populations. He does not, however, derive any generalized formulae. In the present study generalized formulae for both the first moment and the variance of product moments of any order are obtained.

It may be noted here, that all of Pepper's formulae for any infinite population can be obtained from those of the present study as special cases, by assuming that all the populations in the set are identical.

---

[1] A dissertation presented to the Board of Graduate Studies of Washington University in partial fulfilment of the requirements for the degree of Doctor of Philosophy, June 1933.

[2] *Biometrika*, Vol. XXI, Dec. 1929, pp. 231–258.

[3] Church, A. E. R. "On the Means and Squared Standard Deviations of Small Samples from any Population," *Biometrika*, Vol. XVIII, Nov., 1926, pp. 321–394.

[4] *Biometrika*, Vol. XXI, Dec. 1929, pp. 231–258.

[5] Church, A. E. R., "On the Means and Squared Standard Deviations of Small Samples rom any Population," *Biometrika*, Vol. XVIII, Nov., 1926, pp. 321–394.

CHAPTER I. Notations and Definitions

Let (X_1, Y_1), (X_2, Y_2), $\cdots$ (X_n, Y_n) be n bivariate populations each following any law of distribution whatever. The product moment of order a in X and b in Y of the k^{th} population will be denoted by $P^k_{a\,b}$. It is defined as

$$P^k_{a\,b} = E(X_k - a_k)^a\,(Y_k - b_k)^b \tag{1.11}$$

where
$$a_k = E(X_k), \qquad b_k = E(Y_k), \tag{1.12}$$

and where the symbol E signifies the expected value or the mathematical expectation of a quantity.

Regarding each of the n populations of the set as infinite,[6] samples of n are drawn, each member of a sample from one of the n populations.[7] The individual which is drawn from the k^{th} population will be denoted by (x_k, y_k); and the product moment of order a in x and b in y, of such a sample will be denoted by p_{ab}. This product moment may then be defined as

$$p_{ab} = n^{-1}\,S\,(x_k - x)^a\,(y_k - y)^b \tag{1.13}$$

where
$$x = n^{-1}\,Sx_k, \qquad y = n^{-1}\,Sy_k\,. \tag{1.14}$$

The symbols a and b will now be defined by the equations

$$a = n^{-1}\,Sa_k, \qquad b = n^{-1}\,Sb_k\,. \tag{1.15}$$

Obviously
$$E(x) = E(n^{-1}\,Sx_k) = n^{-1}\,SE(X_k) = n^{-1}\,Sa_k = a. \tag{1.16}$$

Similarly $E(y) = b$. That is, the mathematical expectation of the mean, of such a sample as was described above, is equal to the average of the means of all the populations.[8]

In order to make the equations as compact as possible the following additional symbols will be employed:

$$\begin{aligned} x_k - a_k &= u_k, \qquad x - a = u, \qquad \text{and } u_k - u = U_k \\ y_k - b_k &= v_k, \qquad y - b = v, \qquad \text{and } v_k - v = V_k \end{aligned} \tag{1.17}$$

also $a_k - a = A_k$, $b_k - b = B_k$.

From the above definitions it easily follows that

$$E(u_k) = E(v_k) = E(U_k) = E(V_k) = E(u) = E(v) = 0\,. \tag{1.18}$$

---

[6] The term infinite is used here in the probability sense. It is defined very clearly by Church in his "Means and Squared Standard Deviations of Small Samples," *Biometrika*, Vol. XVIII, Nov., 1926, p. 322.

[7] It may be easily shown that this is equivalent to drawing a sample of n from a set of any finite number of populations. The number drawn from each population, however, must be specified. See *Biometrika*, Vol. XIII, 1920–21, p. 295, footnote.

[8] This, of course, is a result of the Lexis Theory, for Poisson and Lexis Series.

The notation is now completed with the definition of the symbol Q_{ij} by the equation:

$$Q_{ij} = S(a_k - a)^i (b_k - b)^j = SA_k^i B_k^j . \tag{1.19}$$

CHAPTER II. The Mathematical Expectation of p_{ab}

The mathematical expectation of p_{ab} will be denoted by $\bar{p}_{ab}$. In the terminology of moments this would be called the mean or first moment of the distribution of p_{ab}.

1. The Mathematical Expectation of p_{11}. According to the above notation the expected value of p_{11} is $\bar{p}_{11}$. By definition

$$\bar{p}_{11} = E(p_{11}) = En^{-1}S(x_i - x)(y_i - y), \tag{2.11}$$

and obviously $En^{-1}S(x_i - x)(y_i - y) = n^{-1}SE(x_i - x)(y_i - y)$.
Writing

$$x_i - x = [(x_i - a_i) - (x - a)] + [a_i - a] = U_i + A_i$$

$$y_i - y = [(y_i - b_i) - (y - b)] + [b_i - b] = V_i + B_i,$$

equation (2.11) may be written as

$$\bar{p}_{11} = n^{-1}SE(U_i + A_i)(V_i + B_i)$$

$$= n^{-1}SE(U_iV_i) + n^{-1}SA_iE(V_i) + n^{-1}SB_iE(U_i) + n^{-1}SE(A_iB_i).$$

Since for any given population A_i and B_i are constants, it follows that $E(A_iB_i) = A_iB_i$. Hence

$$n^{-1}SE(A_iB_i) = n^{-1}SA_iB_i = n^{-1}Q_{11} .$$

Making use of (1.18), it is seen that the terms $n^{-1}SA_iE(V_i)$ and $n^{-1}SB_iE(U_i)$ are zero. The only term left to evaluate is therefore $n^{-1}SE(U_iV_i)$. Since U_i and V_i are symmetric functions of the corresponding small letters, their product is symmetric in u_iv_i. There is therefore no loss in generality if attention is concentrated on a single subscript, say 1.

We may therefore write

$$n^{-1}SE(U_iV_i) = n^{-1}E(U_1V_1) + n^{-1}\underset{*2}{S}E(U_iV_i).$$

Remembering that $U_i = u_i - u = u_i - n^{-1}Su_i$, we may write,

$$U_i = u_i - u = u_i - n^{-1}(u_1 + u_2 + \cdots + u_n)$$

$$= n^{-1}[n_1u_i - (u_1 + u_2 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_n)]$$

---

*The 2 at the bottom of the S simply indicates that the summation begins with $i = 2$.

where $n_1 = n - 1$. In general, n_i will denote the number $n - i$. Similarly

$$V_i = n^{-1}[n_1 v_1 - (v_1 + v_2 + \cdots + v_{i-1} + v_{i+1} + \cdots + v_n)] .$$

Thus

$$n^{-1}SE(U_iV_i) = n^{-3}E(n_1u_1 - u_2 - \cdots - u_n)(n_1v_1 - v_2 - \cdots - v_n)$$

$$+ n^{-3}SE(n_1u_i - u_1 - \cdots - u_{i-1} - u_{i+1} - \cdots u_n)$$

$$(n_1v_i - v_1 - \cdots - v_{i-1} - v_{i+1} - \cdots v_n) .$$

When the right hand side of the last equation is expanded the only terms which appear are of the form $E(u_iv_i)$ and $E(u_iv_j)$. The last one must vanish for u_i and v_j are independent and hence $E(u_iv_j) = E(u_i)E(v_j) = 0$. From the last equation above it is easily seen that the coefficient of $E(u_1v_1)$ is

$$n^{-3}(n_1^2 + n_1) = n^{-3}\, n_1(n_1 + 1) = n^{-2}\, n_1;$$

and because of the symmetry this is obviously the coefficient of any term of that form. Hence

$$n^{-1}SE(U_iV_i) = n^{-2}n_1SE(u_iv_i) .$$

Since $u_i = x_i - a_i$, $v_i = y_i - b_i$, then

$$E(u_iv_i) = E(x_i - a_i)(y_i - b_i) = E(X_i - a_i)(Y_i - b_i) = P^i_{11}$$

and in general,

$$E(u^i_k v^j_k) = P^k_{i\,j} . \tag{2.12}$$

We thus get the formula

$$\tilde{p}_{11} = n^{-2}n_1SP^i_{11} + n^{-1}Q_{11} . \tag{1}$$

Now suppose all the n populations are identical. Then all the A's and also all the B's vanish and therefore, $Q_{11} = 0$. The formula (1) thus becomes

$$\tilde{p}_{11} = \frac{n-1}{n} P_{11} . \tag{1'}$$

This is exactly Pepper's formula for $\tilde{p}_{11}$ for an infinite population.[9]

2. **The Mathematical Expectation of p_{21}.** By definition

$$\tilde{p}_{21} = En^{-1}S(x_i - x)^2(y_i - y) . \tag{2.21}$$

---

[9] *Biometrika*, Vol. XXI, p. 233, Eq. A, N = ∞. As was already stated in the introduction, all the formulae of the present study reduce to Pepper's when the above assumption is made.

Proceeding as above it is seen that

$$En^{-1}S(x_i - x)^2(y_i - y) = n^{-1}SE(x_i - x)^2(y_i - y)$$

$$= n^{-1}SE(U_i + A_i)^2(V_i + B_i) = n^{-1}SE(U_i^2 V_i) + 2n^{-1}SE(U_i V_i A_i)$$

$$+ n^{-1}SE(U_i^2 B_i) + n^{-1}SE(V_i A_i^2) + 2n^{-1}SE(A_i B_i U_i) + n^{-1}SE(A_i^2 B_i) \cdots (2.22)$$

It is quite evident that the two terms before the last vanish. To evaluate the remaining terms, we employ the reasoning of section 1 of this chapter and write:

$$SE(U_i^2 V_i) = E(U_1^2 V_1) + \underset{2}{SE}(U_i^2 V_i)$$

$$= n^{-3}E(n_1^* u_1 - u_2 - \cdots)(n_1 v_1 - v_2 - \cdots) + n^{-3}\underset{2}{SE}(n_1 u_i - u_1 - \cdots)$$

$$(n_1 v_i - v_1 - \cdots).$$

Since terms of the form $E(u_i^2 v_j)$ vanish, only the coefficient of the term $E(u_i^2 v_i)$ must be found. Again considering the subscript 1, the coefficient of $E(u_1^2 v_1)$ is easily found from the last equation to be

$$n^{-3}(n_1^3 - n_1) = n^{-3}n_1(n_1 + 1)(n_1 - 1) = n^{-2}n_1 n_2 .$$

Thus

$$n^{-1}SE(U_i^2\ V_i) = n^{-2}n_1 n_2 SE(u_i^2\ v_i) = n^{-2}n_1 n_2 SP_{2\,1}^{i} . \tag{2.23}$$

For the second term of (2.22) we have

$$SE(U_i V_i A_i) = E(U_1 V_1 A_1) + \underset{2}{SE}(U_i V_i A_i)$$

$$= n^{-2}E(n_1 u_1 - u_2 - \cdots)(n_1 v_1 - v_2 - \cdots)A_1 + n^{-2}\underset{2}{SE}(n_1 u_i - u_1 - \cdots)$$

$$(u_1 v_i - v_1 - \cdots)A_i .$$

The coefficient of $E(u_1 v_1)$ in the first term of the right hand side of the last equation is $n^{-2}n_1^2 A_1$. In the second term it is $n^{-2}\underset{2}{SA_i} = -n^{-2}A_1$, since $SA_i = 0$. It therefore follows that

$$2n^{-1}SE(U_i V_i A_i) = 2n^{-2}n_2 SP_{1\,1}^{i} A_i . \tag{2.24}$$

Quite similarly

$$n^{-1}SE(U_i^2 B_i) = n^{-2}n_2 SP_{2\,0}^{i} B_i , \tag{2.25}$$

and it is obvious that

$$n^{-1}SE(A_i^2 B_i) = n^{-1}Q_{21} . \tag{2.26}$$

---

* Note that the u which has the coefficient n_1 does not occur among the u's which have the negative sign.

We thus get the formula

$$\bar{p}_{21} = n^{-3}n_1n_2SP^i_{21} + n^{-2}n_2S(2P^i_{11}A_i + P^i_{20}B_i) + n^{-1}Q_{21}. \tag{2}$$

3. **The Mathematical Expectation of p_{31} and p_{22}.**

$$\bar{p}_{31} = En^{-1}S(x_i - x)^3(y_i - y) = n^{-1}SE(x_i - x)^3(y_i - y)$$
$$= n^{-1}SE(U_i + A_i)^3(V_i + B_i) = n^{-1}S\{E(U^3_iV_i + U^3_iB_i + 3U^2_iV_iA_i$$
$$+ 3U^2_iA_iB_i + 3U_iV_iA^2_i + 3U_iA^2_iB_i + V_iA^3_i + A^3_iB_i)\}. \tag{2.31}$$

The two terms before the last are zero. The last term is

$$n^{-1}SE(A^3_iB_i) = n^{-1}Q_{31}. \tag{2.32}$$

By (2.23) and (2.24) and some slight manipulation

$$3n^{-1}SE(U^2_iA_iB_i + U_iV_iA^2_i)$$
$$= 3n^{-3}n_2S(P^i_{20}A_iB_i + P^i_{11}A^2_i) + 3n^{-3}(Q_{11}SP^i_{20} + Q^i_{20}SP^i_{11}), \tag{2.33}$$

and by (2.22)

$$n^{-1}SE(U^3_iB_i + 3U^2_iV_iA_i) = n^{-4}(n^3_1 + 1)S(P^i_{30}B_i + 3P^i_{21}A_i). \tag{2.34}$$

The only new term which is to be evaluated is $SE(U^3_iV_i)$. This may be written as follows:

$$SE(U^3_iV_i) = n^{-4}SE(n_1u_i - u_1 - \cdots)^3(n_1v_i - v_1 - \cdots).$$

When the right hand side is expanded it is found that the only non-vanishing terms are of the form $E(U^3_iV_i)$ and $E(u^2_iu_jv_j)$. Only two subscripts, therefore, have to be considered. Without any loss in generality these may be taken as 1 and 2, and the right hand side of the last equation may then be written as follows:

$$SE(n_1u_i - u_1 - \cdots)^3(n_1v_i - v_1 - \cdots) = E(n_1u_1 - u_2 - \cdots)^3(n_1v_1 - v_2 - \cdots)$$
$$+ E(n_1u_2 - u_1 - \cdots)^3(n_1v_1 - v_2 - \cdots) + \underset{3}{SE}(n_1u_i - u_1 - u_2 - \cdots)^3$$
$$\cdot \qquad (n_1v_i - v_1 - v_2 - \cdots).$$

From this last expansion it is easily seen that the coefficient of $E(u^3_iv_i)$ is $(n^4_1 + n_1)$ and that of $E(u^2_iu_jv_j)$, $(6n^2_1 + 3n_2) = 3(2n^2_1 + n_2)$. We thus finally obtain

$$SE(U^3_iV_i) = n^{-4}\{(n^4_1 + n_1)SE(u^3_iv_i) + 3(2n^2_1 + n_2)SE(u^2_iu_jv_j)\}.$$

But by (2.12) $E(u^3_iv_i) = P^i_{31}$, and since u_i and u_j and u_i and v_j are independent $E(u^2_iu_jv_j) = E(u^2_i)E(u_jv_j) = P^i_{20}P^j_{11}$. Whence

$$E(U^3_iV_i) = n^{-4}\{(n^4_1 + n_1)SP^i_{31} + 3(2n^2_1 + n_2)SP^i_{20}P^j_{11}\}. \tag{2.35}$$

From (2.31) and the succeeding equations we finally get

$$\bar{p}_{31} = n^{-5}\{(n_1^4 + n_1)SP^i_{3\,1} + 3(2n_1^2 + n_2)SP^{*i}_{2\,0}P^j_{1\,1}\}$$
$$+ n^{-4}\{(n_1^3 + 1)S(P^i_{3\,0}B_i + 3P^i_{2\,1}A_i)\} + 3n^{-3}\{(n_1^2 - 1)S(P^i_{2\,0}A_iB_i + P^i_{1\,1}A_i^2)$$
$$+ Q_{11}SP^i_{2\,0} + Q_{20}SP^i_{1\,1}\} + n^{-1}Q_{31}. \tag{3}$$

The derivation of $\bar{p}_{22}$ is so similar to that of $\bar{p}_{31}$, that it would be mere repetition to go through the details again. We shall therefore merely write down the formula for $\bar{p}_{22}$ which is

$$\bar{p}_{22} = n^{-5}\{(n_1^4 + n_1)SP^i_{2\,2} + (2n_1^2 + n_2)S(P^i_{2\,0}P^j_{0\,2} + 4P^i_{1\,1}P^j_{1\,1})\}$$
$$+ 2n^{-4}\{(n_1^3 + 1)S(P^i_{2\,1}B_i + P^i_{1\,2}A_i)\} + n^{-3}\{(n_1^2 - 1)S(P^i_{2\,0}B_i^2 + 4P^i_{1\,1}A_iB_i$$
$$+ P^i_{0\,2}A_i^2) + Q_{20}SP^i_{0\,2} + Q_{02}SP^i_{2\,0} + 4Q_{11}SP^i_{1\,1}\} + n^{-1}Q_{22}. \tag{4}$$

4. **The Mathematical Expectation of the General Product Moment** p_{ab}. So far, formulae for the mathematical expectation of p_{ab}, for particular values of a and b, have been derived. The method used in deriving these is, however, perfectly general, and now, that it has been sufficiently illustrated, it can be easily generalized.

By definition we have

$$\bar{p}_{ab} = E[n^{-1}S(x_i - x)^a(y_i - y)^b].$$

Making use of the notation of Chapter I this may be written as

$$n\bar{p}_{ab} = ES(U_i + A_i)^a(V_i + B_i)^b = \sum_{q,r=0}^{a,b} C_q^a C_r^b \sum_1^n SE(U_i^{a-q}V_i^{b-r}A_i^qB_i^r) \tag{2.41}$$

where

$$C_q^a = \frac{a!}{q!(a-q)!}, \qquad C_r^b = \frac{b!}{r!(b-r)!}.$$

Expressing the U's and V's in terms of the u's and v's and setting $a - q = l$, $b - r = m$; we may write for a particular pair of values q and r:

$$n^{l+m}SE(U_i^lV_i^mA_i^qB_i^r) = SE(n_1u_i - u_1 - \cdots)^l(n_1v_i - v_1 - \cdots)^mA_i^qB_i^r. \tag{2.42}$$

Consider, now, the general term in the expansion of the right hand side of (2.42). It is of the form:

$$\frac{l!m!}{\Pi\alpha_h!\Pi\beta_h!}(-1)^{l+m}(-n_1)^{\alpha_h+\beta_h}E(n_1u_{j_1}^{\alpha_1}\cdots u_{j_k}^{\alpha_k}v_{j_1}^{\beta_1}\cdots v_{j_k}^{\beta_k}A_i^qB_i^r), \tag{2.43}$$

where $\Pi\alpha_h! = \alpha_1!\alpha_2!\cdots\alpha_k!$

---

* In this case, and also in the formulae that follow, whenever two or more indices appear in a summation, it will be understood that no two of them can have the same value simultaneously.

For particular sets of values $j_1, j_2, \cdots j_k$, $\alpha_1, \alpha_2, \cdots \alpha_k$, and $\beta_1, \beta_2, \cdots \beta_k$, this term will appear in every member of the summation of the right hand side of (2.42), and its coefficient will differ only in the exponent of $(-n_1)$ and in the subscript i of A^qB^r. Because of the symmetry there is no loss in generality if we take for $j_1, j_2, \cdots j_k$, the first k integers. We now break up the summation of the right hand side of (2.42) as follows:

$$\begin{aligned}
&\overset{n}{S}_{1} E(n_1u_i - u_1 - \cdots)^l(n_1v_i - v_1 - \cdots)^m A_i^q B_i^r \\
&\quad = E(n_1u_1 - u_2 - \cdots)^l(n_1v_1 - v_2 - \cdots)^m A_1^q B_1^r \\
&\quad + E(n_1u_2 - u_1 - \cdots)^l(n_1v_2 - v_1 - \cdots)^m A_2^q B_2^r + \cdots + E(n_1u_k - u_1 - \cdots)^l \\
&\quad (n_1v_k - v_1 - \cdots)^m A_k^q B_k^r + \overset{n}{S}_{i=k+1} E(n_1u_i - u_1 - \cdots)^l \\
&\quad (n_1v_i - v_1 - \cdots)^m A_i^q B_i^r . \qquad (2.44)
\end{aligned}$$

From (2.44) we easily get for the total coefficient (excluding the numerical factor) the expression

$$\overset{k}{S}_{h=1} (-n_1)^{\alpha_h+\beta_h} A_h^q B_h^r + \overset{n}{S}_{h=k+1} A_h^q B_h^r .$$

Writing

$$\overset{n}{S}_{k+1} A_h^q B_h^r = \overset{n}{S}_{1} A_h^q B_h^r - \overset{k}{S}_{1} A_h^q B_h^r = Q_{qr} - \overset{k}{S}_{1} A_h^q B_h^r ,$$

the general term, (2.43), together with the total coefficient, may then be written as

$$(-1)^{l+m} \frac{l\,!\; m\,!}{\Pi\alpha_h\,!\;\Pi\beta_h\,!} \left\{ \overset{k}{S}_{h=1} [(-n_1)^{\alpha_h+\beta_h} - 1]\, A_h^q B_h^r + Q_{qr} \right\} \overset{k}{E}_{h=1} u_h^{\alpha_h} v_h^{\beta_h} .$$

Since u_i and u_j, v_i and v_j, and u_i and v_j are independent while u_i and v_i are not, we have:

I. $E\Pi u_h\, v_h = \Pi E u_h\, v_h = \Pi P^h_{\alpha_h\beta_h}$

II. Any term in which $\alpha_h + \beta_h = 1$ must vanish.

From II it follows that the maximum number of subscripts which can appear in any term in the expansion of (2.42), i.e. the upper limit of k, which will be denoted by t, cannot exceed $(l+m)/2$. In fact when $l+m$ is even, $t = (l+m)/2$, while when $l+m$ is odd, t is the largest integer less than $(l+m)/2$.

Making use of (2.41), the equations following it, and the reasoning of the last paragraph, we finally get the formula:

$$\begin{aligned}
n(-n)^{a+b}\tilde{p}_{ab} = (a\,!)\,(b\,!) \overset{n}{S}_{j_h=1} \overset{a,b}{S}_{q,r=0} \frac{(-n)^{q+r}}{q\,!\;r\,!} \overset{a-q,b-r}{S}_{\alpha_h=0,\beta_h=0} \overset{t}{S}_{k=1} \\
\left\{ \overset{k}{S}_{h=1} [(-n_1)^{\alpha_h+\beta_h} - 1]\, A_{j_h}^q B_{j_h}^r + Q_{qr} \right\} \Pi \frac{P^{j_h}_{\alpha_h\beta_h}}{\alpha_h\,!\;\beta_h\,!} . \qquad (5)
\end{aligned}$$

The following restrictions on the α's and β's must be observed

$$\text{(a)}\quad \alpha_1 + \alpha_2 + \cdots + \alpha_k = a - q$$

$$\text{(b)}\quad \beta_1 + \beta_2 + \cdots + \beta_k = b - r$$

$$\text{(c)}\quad \alpha_h + \beta_h \neq 1\,.$$

In case the n populations are identical (5) reduces as follows: For $q = 0$, $r = 0$, $A_i^0 = 1$, $B_i^0 = 1$, and $Q_{00} = n$; while in every other case $A_i^q B_i^r = 0$, $Q_{qr} = 0$. The summations with respect to q and r, therefore disappear.

Consider now the summations

$$\sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_k=1}^{n} P^{j_1} P^{j_2} \cdots P^{j_k}\,.$$

Since all the populations are the same we may drop the j by actually carrying out the indicated summations. If, then, there are c repetitions among the k pairs of integers $\alpha_h\beta_h$, in which $\alpha_1\beta_1$, $\alpha_2\beta_2$, $\cdots$ $\alpha_c\beta_c$, are repeated l_1, l_2, $\cdots$ l_c times respectively, then we have;

$$\sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} P^{j_h}_{\alpha_h\beta_h} = \frac{k!\, C_k^n *}{l_1!\, l_2! \cdots l_c!} \Pi\, P_{\alpha_h\beta_h}\,.$$

We thus arrive at the following corollary: The mathematical expectation, $\bar{p}_{ab}$, of the product moment, p_{ab}, in samples of n from a single infinite population having any law of distribution is given by

$$n(-n)^{a+b}\bar{p}_{ab} = \sum_{\alpha_h,\beta_h=0}^{a,b} \frac{(a!)\,(b!)}{\Pi\alpha_h!\;\Pi\beta_h!} \sum_{k=1}^{t} \left[\sum_{h=1}^{n} (-n_1)^{\alpha_h+\beta_h} + n_k \right] \frac{k!\, C_k^n}{l_1! \cdots l_c!} \Pi\, P^*_{\alpha_h\beta_h}\,. \quad (5')$$

Note: In deriving these general formulae it was assumed that $n > t$. There is however, no loss in generality in this assumption. For, if $t > n$, we may suppose that, $x_{n+1} = x_{n+2} = \cdots = x_c = 0$, and hence $P^{n+1}_{\alpha\beta} = \cdots = P^t_{\mu\nu} = 0$, and thus the above reasoning is still valid.

5. **Formulae for $\bar{p}_{41}$, $\bar{p}_{32}$, $\bar{p}_{51}$, $\bar{p}_{42}$, $\bar{p}_{33}$.** Formulae for $\bar{p}_{ab}$ in which $a + b = 5, 6, 7, 8$ have been obtained. But for $(a + b) > 6$ these formulae become very long, and since these will be of no use in the subsequent work, only those of order 5 and 6 are given below.

$$\bar{p}_{41} = n^{-6}\{(n_1^5 - n_1)SP^i_{41} + 2nn_2^2 S(2P^i_{30}P^j_{11} + 3P^i_{21}P^j_{20})\}$$

$$+ n^{-5}\{(n_1^4 + n_1)S(P^i_{40}B_i + 4P^i_{31}A_i) + 6nn_2 S(P^j_{20}B_iP^j_{20} + 2P^{i\dagger}_{11}A_iP^j_{20})\}$$

---

* This is a generalization of Pepper's results for $N = \infty$. See *Biometrika* Vol. XXI, pp. 231–240.

† The symbol $P^i_{11}A_iP^j_{20}$ is an abbreviation of the full term $(A_i + A_j)\,(P^i_{11}P^j_{20} + P^j_{11}P^i_{20})$. Similar abbreviations will be used in the other formulae.

$$+ 2n^{-4}\{(n_1^3+1)S(2P_{30}^iA_iB_i + 3P_{21}^iA_i^2) - 2Q_{11}SP_{30}^i - 3Q_{20}SP_{21}^i\}$$
$$+ 2n^{-3}\{(n_1^2-1)S(2P_{11}^iA_i^3 3P_{20}^iA_i^2B_i) + 2Q_{30}SP_{11}^i + 3Q_{21}SP_{20}^i\} + n^{-1}Q_{41}. \quad (6)$$

$$\bar{p}_{32} = n^{-6}\{(n_1^5 - n_1)SP_{32}^i + nn_2^2S(P_{30}^iP_{02}^j + 6P_{21}^iP_{11}^j + 3P_{20}^iP_{12}^j)\}$$
$$+ n^{-5}\{(n_1^4-1)S(2P_{31}^iB_i + 3P_{22}^iA_i) + 3nn_2S(P_{20}^iP_{11}^jB_i + [P_{20}^iP_{02}^j$$
$$+ 4P_{11}^iP_{11}^j]A_i)\} + n^{-4}\{(n_1^3+1)S(P_{30}^iB_i^2 + 6P_{21}^iA_iB_i$$
$$+ 3P_{12}^iA_i^2) - Q_{02}SP_{30}^i - 6Q_{11}SP_{21}^i - 3Q_{20}SP_{12}^i\} + n^{-3}\{(n_1^2-1)S(3P_{20}^iA_iB_i^2$$
$$+ 6P_{11}A_iB_i + P_{02}^iA_i^3) + 3Q_{12}SP_{20}^i + 6Q_{21}SP_{11}^i + Q_{30}SP_{02}^i\} + n^{-1}Q_{32}. \quad (7)$$

$$\bar{p}_{51} = n^{-7}\{(n_1^6+n_1)SP_{51}^i + 5(n_1^4+n_1^2+n_2)S(P_{40}^iP_{11}^j$$
$$+ 2P_{31}^iP_{20}^j) - 10(2n_1^3-n_2)SP_{21}^iP_{30}^j + 30(3n_1^2+n_3)SP_{20}^iP_{20}^jP_{11}^k\} + n^{-6}\{(n_1^5$$
$$+ 1)S(P_{50}^iB_i + 5P_{41}^iA_i) + 10(n_1^3+1)S^*[2P_{30}^iP_{20}^jB_i + (2P_{30}^iP_{11}^j$$
$$+ 3P_{21}^iP_{20}^j)A_i] - 10nn_2S^*[2P_{30}^iP_{20}^jB_j + (2P_{30}^iP_{11}^j + 3P_{21}^iP_{20}^j)A_j]\}$$
$$+ 5n^{-5}\{(n_1^4-1)S(P_{40}^iA_iB_i + 2P_{31}^iA_i^2) + 6nn_2S(P_{20}^iP_{20}^jA_iB_i + 2P_{20}^iP_{11}^jA_i^2)$$
$$+ Q_{11}S(P_{40}^i + 6P_{20}^iP_{20}^j) + 2Q_{20}S(P_{31}^i + 6P_{20}^iP_{11}^j)\} + 10n^{-4}\{(n_1^3$$
$$+ 1)S(P_{30}^iA_i^2B_i + P_{21}^iA_i^3) - Q_{21}SP_{30}^i - Q_{30}SP_{21}^i\} + 5n^{-3}\{(n_1^3-1)S(2P_{20}^iA_i^3B_i$$
$$+ P_{11}^iA_i^4) + 2Q_{31}SP_{20}^i + Q_{40}SP_{11}^i\} + n^{-1}Q_{51}. \quad (8)$$

$$\bar{p}_{42} = n^{-7}\{(n_1^6+n_1)SP_{42}^i + (n_1^4+n_1^2+n_2)S(P_{40}^iP_{02}^j + 8P_{31}^iP_{11}^j$$
$$+ 6P_{22}^iP_{20}^j) + 4(2n_1^3-n_2)S(P_{30}^iP_{12}^j + 3P_{21}^iP_{21}^j) + 6(3n_1^2+n_3)S(P_{20}^iP_{20}^jP_{02}^k$$
$$+ 4P_{20}^iP_{11}^jP_{11}^k)\} + 2n^{-6}\{(n_1^5+1)S(P_{41}^iB_i + 2P_{32}^iA_i) + 2(n_1^3+1)S[(2P_{30}^iP_{11}^j$$
$$+ 3P_{21}^iP_{20}^j)B_i + (P_{30}^iP_{02}^j + 6P_{21}^iP_{11}^j + 3P_{12}^iP_{20}^j)A_i] - 2nn_2S[(2P_{30}^iP_{11}^j$$
$$+ 3P_{21}^iP_{20}^j)B_j + (P_{30}^iP_{12}^j + 6P_{21}^iP_{11}^j + 3P_{12}^iP_{20}^j)A_j]\} + n^{-5}\{(n_1^4-1)S(P_{40}^iB_i^2$$
$$+ 8P_{31}^iA_iB_i + 6P_{22}^iA_i^2) + 6nn_2S[P_{20}^iP_{21}^jB_i^2 + 4P_{20}^iP_{11}^iA_iB_i + (P_{20}^iP_{02}^j$$
$$+ 4P_{11}^iP_{11}^j)A_i^2] + Q_{02}S(P_{40}^i + 6P_{20}^iP_{20}^j) + 8Q_{11}S(P_{31}^i + 3P_{20}^iP_{11}^j)$$
$$+ 6P_{20}S(P_{22}^i + P_{20}^iP_{02}^j + 4P_{11}^iP_{11}^j)\} + 4n^{-4}\{(n_1^3+1)S(P_{30}^iA_iB_i^2$$
$$+ 3P_{21}^iA_i^2B_i + P_{12}^iA_i^3) - Q_{12}SP_{30}^i - 3Q_{21}SP_{21}^i + Q_{30}^iSP_{12}^i\}$$
$$+ n^{-3}\{S(6P_{20}^iA_i^2B_i^2 + 8P_{11}^iA_i^3B_i + P_{02}^iA_i^4) + Q_{40}SP_{02}^i + 8Q_{31}SP_{11}^i$$
$$+ 6Q_{22}SP_{20}^i\} + n^{-1}Q_{42}. \quad (9)$$

$$\bar{p}_{33} = n^{-7}\{(n_1^6+n_1)SP_{33}^i + 3(n_1^4+n_1^2+n_2)S(P_{31}^iP_{p2}^j + 3P_{22}^iP_{11}^j$$
$$+ P_{13}^iP_{20}^j) - (2n_1^3-n_2)S(P_{30}^iP_{03}^j + 9P_{21}^iP_{12}^j) + 9(3n_1^3+n_3)S(P_{20}^iP_{11}^jP_{02}^k$$

---

* The repetition of this expression signifies that A and B factors are coupled only with those P factors which have corresponding indices.

$$+ 4P^i_{11}P^j_{11}P^k_{11})\} + 3n^{-6}\{(n_1^5 + 1)S(P^i_{32}B_i + P^i_{23}A_i) + (n_1^3 + 1)\, S[(P^i_{30}P^j_{02}$$
$$+ 6P^i_{21}P^j_{11} + 3P^i_{12}P^j_{20})B_i + (P^i_{03}P^j_{20} + 6P^i_{12}P^j_{11} + 3P^i_{21}P^j_{02})A_i]$$
$$- nn_2S[(P^i_{30}P^j_{02} + 6P^i_{21}P^j_{11} + 3P^i_{12}P^j_{20})B_j + (P^i_{03}P^j_{20} + 6P^i_{12}P^j_{11}$$
$$+ 3P^i_{21}P^j_{02})A_j]\} + 3n^{-5}\{(n_1^4 - 1)S(P^i_{31}B_i^2 + 3P^i_{22}A_iB_i + P^i_{13}A_i^2)$$
$$+ 3n_1n_2S[P^i_{20}P^j_{11}B_i^2 + (P^i_{20}P^j_{02} + 4P^i_{11}P^j_{11})A_iB_i + P^i_{11}P^j_{02}A_i^2)]$$
$$+ S[Q_{02}(P^i_{31} + 3P^i_{20}P^j_{11}) + 3Q_{11}(P^i_{22} + P^i_{20}P^j_{02} + 4P^i_{11}P^j_{11}) + Q_{20}(P^i_{13}$$
$$+ 3P^i_{02}P^i_{11})]\} + n^{-4}\{(n_1^3 + 1)S(P^i_{30}B_i^3 + 9P^i_{21}A_iB_i^2 + 9P^i_{12}A_i^2A_i + P^i_{03}A_i^3)$$
$$- S(Q_{03}P^i_{30} + 9Q_{12}P^i_{21} + 9Q_{21}P^i_{12} + Q_{30}P^i_{03})\} + 3n^{-3}\{(n_1^2 - 1)S(P^i_{20}A_iB_i^3$$
$$+ 3P^i_{11}A_i^2B_i^2 + P^i_{02}A_i^3B_i) + S(Q_{13}P^i_{20} + 3Q_{22}P^i_{11} + Q_{31}P^i_{02})\} + n^{-1}Q_{33}. \quad (10)$$

CHAPTER III. The Mathematical Expectation of the Variance of p_{ab}

1. The Symbols $_2m_{p_{ab}}$ and $_2M_{p_{ab}}$. Denoting the variance of p_{ab} by m and the mathematical expectation of $_2m_{p_{ab}}$ by $_2M_{p_{ab}}$, we have the definition,

$$_2m_{p_{ab}} = \{n^{-1}S(x_i - x)^a(y_i - y)^b - \bar{p}_{ab}\}^2$$
$$= n^{-2}S^2(x_i - x)^a(y_i - y)^b - 2n^{-1}\bar{p}_{ab}S(x_i - x)^a(y_i - y)^b + \bar{p}_{ab}^2, \text{ and}$$
$$_2M_{p_{ab}} = E(_2m_{p_{ab}}) = E\{n^{-2}S^2(x_i - x)^a(y_i - y)^b - 2n^{-1}\bar{p}_{ab}S(x_i - x)^a(y_i - y)^b + \bar{p}_{ab}^2\}$$
$$= n^{-2}E[S(x_i - x)^{2a}(y_i - y)^{2b}] + 2n^{-2}E[S(x_i - x)^a(x_j - x)^a(y_i - y)^b(y_j - y)^b]$$
$$- 2n^{-1}\bar{p}_{ab}E[S(x_i - x)^a(y_i - y)^b] + \bar{p}_{a\,b}^2 = n^{-1}\bar{p}_{2a2b}$$
$$+ 2n^{-2}E[S(x_i - x)^a(y_i - y)^b(x_j - x)^a(y_j - y)^b] - \bar{p}_{ab}^2. \quad (3.11)$$

Before attempting to expand the right hand side of (3.11) for any values a, b we shall derive the formula for $_2M_{p_{11}}$ to illustrate the procedure.

2. The Mathematical Expectation of $_2m_{p_{11}}$. By (3.11) we have

$$_2M_{p_{11}} = n^{-1}\bar{p}_{22} + 2n^{-2}E[S(x_i - x)(y_i - y)(x_j - x)(y_j - y)] - \bar{p}_{1\,1}^2. \quad (3.21)$$

The first term is given by (4) and the last by (1). The only new term is the middle one. To expand it let us write it in terms of U and V. We then have:

$$n^{-2}SE[(x_i - x)(y_i - y)(x_j - x)(y_j - y)] = n^{-2}SE[(U_i + A_i)(V_i$$
$$+ B_i)(U_j + A_j)(V_j + B_j)] = n^{-2}\{SE[U_iV_iU_jV_j + (U_iV_iU_jB_j + U_jV_jU_iB_i)$$
$$+ (U_iV_iV_jA_j + U_jV_jV_iA_i) + (U_iV_iA_jB_j + U_jV_jA_iB_i)$$
$$+ (U_iV_jA_jB_i + U_jV_iA_iB_j) + U_iU_jB_iB_j + V_iV_jA_iA_j + 4 \text{ vanishing terms}$$
$$+ A_iB_iA_jB_j]\}. \quad (3.22)$$

The evaluation of the last term is very simple. For

$$SE(A_iB_iA_jB_j) = S(A_iB_iA_jB_j),$$

and from the elementary theory of symmetric functions we have:

$$S(A_iB_iA_jB_j) = \frac{S^2(A_iB_i) - S(A_i^2B_i^2)}{2}.$$

Hence

$$SE(A_iB_iA_jB_j) = \frac{S^2(A_iB_i) - S(A_i^2B_i^2)}{2} = \frac{Q_{11}^2 - Q_{22}}{2}. \tag{3.23}$$

To expand the first term and also the remaining ones, we return to the u, v, notation defined in Chapter I. We then write

$$SE(U_iV_iU_jV_j) = n^{-4}SE[(n_1u_i - u_1 - \cdots)(n_1v_i - v_1 - \cdots)$$
$$(n_1u_j - u_1 - \cdots)(n_1v_j - v_1 - \cdots)].$$

The only terms which can appear in the expansion of the right hand side of the last equation have the following form:

$$E(u_i^2v_i^2), \quad E(u_i^2v_j^2), \quad E(u_iv_iu_jv_j),$$

i.e., exactly those which appear in the evaluation of $\bar{p}_{22}$. Remembering the symmetry, there will be no loss in generality if we take for i and j the integers 1 and 2. To find the coefficients of the three characterstic terms, the above summation may be broken up as follows:

$$n^4SE(U_iV_iU_jV_j) = E[(n_1u_1 - u_2 - \cdots)(n_1v_1 - v_2 - \cdots)(n_1u_2 - u_1 - \cdots)$$
$$(n_1v_2 - v_1 - \cdots)] + E\{[n_1u_1 - u_2 - \cdots)(n_1v_1 - v_2 - \cdots) + (n_1u_2 - u_1$$
$$- \cdots)(n_1v_2 - v_1 - \cdots)]\underset{3}{S}(n_1u_i - u_1 - \cdots)(n_1v_i - v_1 - \cdots)\} + \underset{3}{SE}[(n_1u_i$$
$$- u_1 - \cdots)(n_1v_i - v_1 - \cdots)(n_1u_j - u_1 - \cdots)(n_1v_j - v_1 - \cdots)]. \tag{3.24}$$

Writing the three terms in a row and their coefficients from the three parts of (3.24) in columns below these terms, we get the following scheme:

	$E(u_1^2v_1^2)$	$E(u_1^2v_2^2 + u_2^2v_1^2)$	$E(u_1v_1u_2v_2)$
	n_1^2	n_1^2	$(n_1^2 + 1)^2$
	$n_2(n_1^2 + 1)$	$-2n_1n_2$	$2n_2^3$
	$\frac{n_2n_3}{2}$	$\frac{n_2n_3}{2}$	$2n_2n_3$
Total coeff.	$\frac{nn_1(2n_1 - 1)}{2}$	$\frac{-nn_3}{2}$	$n(n_1^3 + n_1^2 - 3n_1 + 3)$.

With the aid of the above equations we finally get:

$$SE(U_iV_iU_jV_j) = n^{-4}\left\{\frac{n_1n(2n_1-1)}{2}\overset{*}{SP^i_{22}} - \frac{nn_3}{2}SP^i_{20}P^j_{02} + n(nn_1^2-3n_2)SP^i_{11}P^j_{11}\right\}$$

Proceeding in the same way we find:

$$SE(U_iV_iU_jB_j + U_jV_jU_iB_i) = n^{-3}(2n_1^2+n_2)SP^i_{21}B_i$$

$$SE(U_iV_iV_jA_j + U_jV_jV_iA_i) = n^{-3}(2n_1^2+n_2)SP^i_{12}A_i$$

$$SE(U_iV_iA_jB_j + U_jV_jA_iB_i) = -nn_2SP^i_{11}A_iB_i + (n_1^2+n_2)Q_{11}SP^i_{11}$$

$$SE(U_iV_jA_jB_i + U_jV_iA_iB_j) = 2nSP^i_{11} - Q_{11}SP^i_{11}$$

$$SE(U_iU_jB_iB_j + V_iV_jA_iA_j) = nS(P^i_{20}B_i^2 + P^i_{02}A_i^2) - \tfrac{1}{2}S(Q_{20}P^i_{02} + Q_{02}P^i_{20}).$$

Collecting terms and simplifying we finally get:

$$_2M_{p_{11}} = n^{-4}\{n_1^2SP^i_{22} + S(P^i_{20}P^j_{02} + 2P^i_{11}P^j_{11}) - n^2S(P^i_{11})^2\}$$
$$+ 2n^{-3}n_1\{S(P^i_{21}B_i + P^i_{12}A_i)\} + n^{-2}\{S(P^i_{20}B_i^2 + 2P^i_{11}A_iB_i + P^i_{02}A_i^2)\}. \quad (11)$$

Corollary 1. In case $X_i = Y_i$, i.e., when the set of populations are univariate, (11) becomes

$$_2M_{p_{20}} = n^{-4}\{n_1^2S[P^i_{40} - (P^i_{20})^2] + 4SP^i_{20}P^j_{20}\} + 4n^{-3}n_1SP^i_{30}A_i + 4n^{-2}SP^i_{20}A_i^2. \quad (11')$$

This is Tchouproff's formula for the expected value of the variance of samples of n.[10]

Corollary 2. In case the n populations are identical (11) becomes

$$_2M_{p_{11}} = n^{-3}n_1[n_1P_{22} + P_{20}P_{02} - n_2P_{11}^2]. \quad (11'')[11]$$

3. **The Mathematical Expectation of $_2M_{p_{ab}}$.** We now return to the general equation

$$_2M_{p_{ab}} = n^{-1}\tilde{p}_{2a2b} - \tilde{p}^2_{ab} + 2n^{-2}\overset{n}{\underset{i=1,j=1}{S}}E(x_i-x)^a(y_i-y)^b(x_j-x)^a(y_j-y)^b. \quad (3.11)$$

---

* Since $E(u_i^2v_i^2) = P^i_{22}$, $E(u_i^2v_j^2) = P^i_{20}P^j_{02}$, etc.

[10] See *Biometrika*, Vol. XIII p. 295.

[11] See *Biometrika*, Vol. XXI p. 234, Cor. 1.

The first two terms are given by (5). To evaluate the last term we write:

$$SE[(x_i - x)^a(y_i - y)^b(x_j - x)^a(y_j - y)^b] = SE[(U_i + A_i)^a(V_i + B_i)^b(U_j + A_j)^a$$

$$(V_j + B_j)^b] = SE(U_i^a V_i^b U_j^a V_j^b) + \underset{r_1, r_2, r_3, r_4 = 0}{S}{}^{a, a, b, b,} C_{r_1}^a C_{r_2}^a C_{r_3}^b C_{r_4}^b$$

$$\overset{n}{\underset{1}{SE}}(U_i^\alpha V_i^\gamma U_j^\beta V_j^\delta A_i^{r_1} B_i^{r_3} A_j^{r_2} B_j^{r_4}) = n^{-2(a+b)} SE\{(n_1 u_i - \cdots)^a (n_1 v_i - \cdots)^b$$

$$(n_1 u_j - \cdots)^a (n_1 v_j - \cdots)^b + \underset{r_1 \cdots r_4}{S} n^{(r_1+r_2+r_3+r_4)} C_{r_1}^a \cdots C_{r_4}^b SE[(n_1 u_i - \cdots)^\alpha$$

$$(n_1 v_i - \cdots)^\gamma (n_1 u_j - \cdots)^\beta (n_1 v_j - \cdots)^\delta A_i^{r_1} B_i^{r_3} A_j^{r_2} B_j^{r_4}, \tag{3.31}$$

where $\alpha = a - r_1, \beta = a - r_2, \gamma = b - r_3, \delta = b - r_4$.

The right hand side of (3.31) has been broken up into two parts because the first part is symmetrical, while the second part, in general, is not except when $r_1 = r_2$, and $r_3 = r_4$.

Let us now consider the expression

$$SE[(n_1 u_i - \cdots)^a (n_1 v_i - \cdots)^b (n_1 u_j - \cdots)^a (n_1 v_j - \cdots)^b]. \tag{3.32}$$

This is a double summation in which $c_{ij} = c_{ji}$ and in which the diagonal terms, c_{ii}, are missing.

Consider next a general term of k factors from the expansion of each bracket of (3.32). As we are dealing with symmetric functions, there will be no loss in generality if we consider the first k subscripts only; and if we let the lower limits of the exponents of the u's and v's begin with zero we may consider that each parenthesis of a given bracket contributes exactly k factors. Such a term, omitting the coefficient, may be written as follows:

$$E(u_1^{\alpha_1} \cdots u_k^{\alpha_k} v_1^{\beta_1} \cdots v_k^{\beta_k} u_1^{\alpha_1'} \cdots u_k^{\alpha_k'} v_1^{\beta_1'} \cdots v_k^{\beta_k'}) = \prod_{h=1}^{k} E(u_h^{\alpha_h + \alpha_h'} v_h^{\beta_h + \beta_h'})$$

$$= \prod_{h=1}^{k} P^h(\alpha_h + \alpha_h')\,(\beta_h + \beta_h'). \tag{3.33}$$

This term occurs in every one of the $\frac{1}{2}nn_1$ brackets of (3.32), having the same numerical coefficient in every one of them, which is

$$\frac{(a\,!)^2\,(b\,!)^2}{\Pi\alpha_h\,!\,\Pi\alpha_h'\,!\,\Pi\beta_h\,!\,\Pi\beta_h'\,!}. \tag{3.34}$$

To obtain the n_1 coefficient of (3.33) we break up (3.32) into the following partial summations:

$$E[(n_1 u_i - \cdots)^a (n_1 v_i - \cdots)^b (n_1 u_j - \cdots)^a (n_1 v_j - \cdots)^b] = E[(n_1 u_1 - \cdots)^a$$

$$(n_1 v_1 - \cdots)^b\,(n_1 u_2 - \cdots)^a (n_1 v_2 - \cdots)^b] + \cdots + E[(n_1 u_{k-1} - \cdots)^a$$

$$(n_1v_{k-1} - \cdots)^b\ (n_1u_k - \cdots)^a(n_1v_k - \cdots)^b] + \sum_{i=1}^{k} E\Big[(n_1u_i - \cdots)^a$$

$$(n_1v_i - \cdots)^b \sum_{j=k+1}^{n} (n_1u_j - \cdots)^a(n_1v_j - \cdots)^b\Big] + \sum_{i,j=k+1}^{n} [E\{(n_1u_i - \cdots)^a$$

$$(n_1v_i - \cdots)^b(n_1u_j - \cdots)^a(n_1v_j - \cdots)^b\}] .$$

From this equation we get for the total coefficient in n of the term (3.33) the following expression:

$$\sum_{h,h'=1}^{k} (-n_1)^{\alpha_h+\alpha_{h'}'+\beta_h+\beta_{h'}'} + n_k \sum_{h=1}^{k} [(-n_1)^{\alpha_h+\beta_h} + (-n_1)^{\alpha_h'+\beta_h'}] + C_2^{n_k} .$$

The following restrictions on the α's and β's must be observed.

$$\text{(a)}\quad \begin{matrix}\alpha_1 + \alpha_2 + \cdots + \alpha_k = a \\ \alpha_1' + \alpha_2' + \cdots + \alpha_k' = a\end{matrix} \qquad \text{(b)}\quad \begin{matrix}\beta_1 + \beta_2 + \cdots + \beta_h = b \\ \beta_1' + \beta_2' + \cdots + \beta_k' = b\end{matrix}$$

$$\text{(c)}\quad \alpha_h + \alpha_h' + \beta_h + \beta_h' \neq 1 .$$

From (c) we obtain the upper limit of k, namely: $t = a + b$.

Combining the various above equations we finally obtain:

$$(n)^{2(a+b)}\, S(U_i^a V_i^b U_j^a V_j^b) = (a!)^2(b!)^2 \sum_{i_h=1}^{n} \sum_{\alpha_h,\alpha_{h'}',\beta_h,\beta_{h'}'=0}^{n,b}$$

$$\sum_{k=1}^{t} \left\{ \sum_{h,h'=1}^{k} (-n_1)^{\alpha_h+\beta_h+\alpha_{h'}'+\beta_{h'}'} + n_k \sum_{h=1}^{k} [(-n_1)^{\alpha_h+\beta_h} + (-n_1)^{\alpha_h'+\beta_h'}] + C_2^{n_k} \right\}$$

$$\frac{\Pi\, P^{J_h}_{(\alpha_h+\alpha_h')(\beta_h+\beta_h')}}{\Pi\alpha_h!\ \Pi\alpha_h'!\ \Pi\beta_h!\ \Pi\beta_h'!} . \tag{3.35}$$

Turning to the second part of (3.31) let us consider the expression

$$\sum_{i=1,j=1}^{n} E[(n_1u_i - \cdots)\,(n_1v_i - \cdots)\,(n_1u_j - \cdots)\,(n_1v_j - \cdots)\, A_i^{r_1}B_i^{r_3}A_j^{r_2}B_j^{r_4}$$

for a given set of r's. The term (3.33) may also be considered as a general term of this last expression; of course, the exponents of the u's and v's will be different in this case. In order to evaluate the complete coefficient of a term like (3.33) we again write;

$$SE[(n_1u_i - \cdots)^\alpha(n_1v_i - \cdots)^\gamma(n_1u_j - \cdots)^\beta(n_1v_j - \cdots)^\delta\, A_i^{r_1}B_i^{r_3}A_j^{r_2}B_j^{r_4}$$

$$= E[(n_1u_1 - \cdots)^\alpha(n_1v_1 - \cdots)^\gamma(n_1u_2 - \cdots)^\beta(n_1v_2 - \cdots)^\delta A_1^{r_1}B_1^{r_3}A_2^{r_2}B_2^{r_4}]$$

$$+ E[n_1u_2 - \cdots)^\alpha(n_1v_2 - \cdots)^\gamma(n_1u_1 - \cdots)^\beta(n_1v_1 - \cdots)^\delta A_2^{r_1}B_2^{r_3}A_1^{r_2}B_1^{r_4}$$

$$+ \cdots + E[(n_1u_k - \cdots)^\alpha(n_1v_k - \cdots)^\gamma(n_1u_{k-1} - \cdots)^\beta(n_1v_{k-1} - \cdots)^\delta$$

$$A_k^{r1}B_k^{r3}A_{k-1}^{r2}B_{k-1}^{r4}] + \sum_{i=1}^{k} E[(n_1u_i - \cdots)^{\alpha}(n_1v_i - \cdots)^{\gamma}A_i^{r1}B_i^{r3} \sum_{j=k+1}^{n}$$

$$(n_1u_j - \cdots)^{\beta}(n_1v_j - \cdots)^{\delta}A_j^{r2}B_j^{r4}] + \sum_{j=1}^{k} E[(n_1u_j - \cdots)^{\beta}(n_1v_j - \cdots)^{\delta}$$

$$A_j^{r2}B_j^{r4} \sum_{i=k+1}^{n} (n_1u_i - \cdots)^{\alpha}(n_1v_i - \cdots)^{\gamma}A_i^{r1}B_i^{r3}] + \sum_{i,j=k+1}^{n} E[(n_1u_i - \cdots)^{\alpha}$$

$$n_1v_i - \cdots)^{\gamma}(n_1u_j - \cdots)^{\beta}(n_1v_j - \cdots)^{\delta}A_i^{r1}B_i^{r3}A_j^{r2}B_j^{r4}]. \tag{3.36}$$

It is now quite easy to write down the complete coefficient of a term of the form (3.33). The numerical coefficient of this term is the same in every bracket of (3.36), and is

$$\frac{(-1)\overset{4}{S}_{1,i}(a-r_1)!\,(a-r_2)!\,(b-r_3)!\,(b-r_4)!}{\Pi\alpha_h!\,\Pi\alpha_h'!\,\Pi\beta_h!\,\Pi\beta_h'!}. \tag{3.37}$$

The coefficient in n_1 and $A_i^{r1}\,B_i^{r3}\,A_j^{r2}\,B_j^{r4}$ is broken up by (3.36) into the following four parts:

I. $\sum_{h=1,h'=1}^{k} (-n_1)^{\alpha_h+\alpha_{h'}'+\beta_h+\beta_{h'}'}A_h^{r1}B_h^{r3}A_{h'}^{r2}B_{h'}^{r4}$, from the first $k(k-1)$ brackets.

II. $\sum_{h=1}^{k} (-n_1)^{\alpha_h+\beta_h}A_h^{r1}B_h^{r3} \sum_{h'=k+1}^{n} A_{h'}^{r2}B_{h'}^{r4} = \sum_{h=1}^{k} (-n_1)^{\alpha_h+\beta_h}A_h^{r1}B_h^{r3}$

$$\left[Q_{r2r4} - \sum_{h'=1}^{k} A_{h'}^{r2}B_{h'}^{r4}\right],$$

from the next $k(n-k)$ brackets. Similarly

III. $\sum_{h'=1}^{k} (-n_1)^{\alpha_{h'}'+\beta_{h'}'}A_{h'}^{r2}B_{h'}^{r4}\left[Q_{r1r3} - \sum_{h=1}^{k} A_h^{r1}B_h^{r3}\right]$, from the next $k(n-k)$.

And finally:

IV. $\sum_{i,j=k+1}^{n} A_i^{r1}B_i^{r3}A_j^{r2}B_j^{r4} = \overset{n}{S}_1 A_h^{r1}B_h^{r3}\overset{n}{S}_1 A_h^{r2}B_h^{r4} - \overset{n}{S}_1 A_h^{(r1+r2)}B_h^{(r3+r4)}$

$$- \sum_{h,h'=1}^{k} A_h^{r1}B_h^{r3}A_{h'}^{r2}B_{h'}^{r4} - \sum_{h=1}^{n} A_h^{r1}B_h^{r3} \sum_{h'=1}^{k} A_{h'}^{r2}B_{h'}^{r4} - \sum_{h=1}^{n} A_h^{r2}B_h^{r4} \sum_{h=1}^{k} A_h^{r1}B_h^{r3}$$

$$+ 2\sum_{h=1}^{k} A_h^{r1}B_h^{r3} \sum_{h'=1}^{k} A_{h'}^{r2}B_{h'}^{r4} = Q_{r1r3}Q_{r2r4} - Q_{(r1+r2)(r3+r4)} - Q_{r1r3}\overset{k}{S}_1 A_h^{r2}B_h^{r4}$$

$$- Q_{r2r4}\overset{k}{S}_1 A_h^{r1}B_h^{r3} - \sum_{h,h'=1}^{k} A_h^{r1}B_h^{r3}A_{h'}^{r2}B_{h'}^{r4} + 2\sum_{h=1}^{k} A_h^{r1}B_h^{r3} \sum_{h'=1}^{k} A_{h'}^{r2}B_{h'}^{r4}$$, from the

last c_2^{nk} brackets.

The restrictions on the α's and β's differ from those given above in that a is replaced by $a - r_1$ and $a - r_2$, and b by $b - r_3$ and $b - r_4$; and from the restriction (c) we get for the upper limit of k, in this case,

$$t_1 = \frac{\alpha + \beta + \gamma + \delta}{2} = a + b - \frac{r_1 + r_2 + r_3 + r_4}{2}$$

when $\overset{4}{\underset{1}{S}} r_i$ is even, or the greatest integer less then $\frac{S\alpha}{2}$ when Sr_i is odd.

Combining (3.37) with $C^a_{r_1} \cdots C^b_{r_4}$ we get for the general numerical coefficient in the expansion of the second part of (3.31), the expression

$$\frac{(-1)Sr_i\,(a!)^2\,(b!)^2}{\Pi r_i!\,\Pi\alpha_h!\,\Pi\alpha_h'!\,\Pi\beta_h!\,\Pi\beta_h'!}.$$

By an obvious manipulation we have

$$\begin{aligned}
\mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} = {} & \overset{k}{\underset{h,h'=1}{S}} \left[(-n_1)^{\alpha_h+\beta_h+\alpha_{h'}'+\beta_{h'}'} - 1\right] A_h^{r_1} B_h^{r_3} A_{h'}^{r_2} B_{h'}^{r_4} + Q_{r_2 r_4} \\
& \overset{k}{\underset{h=1}{S}} \left[(-n_1)^{\alpha_h+\beta_h} - 1\right] A_h^{r_1} B_h^{r_3} + Q_{r_1 r_3} \overset{k}{\underset{h=1}{S}} \left[(-n_1)^{\alpha_h'+\beta_h'} - 1\right] A_h^{r_2} B_h^{r_4} \\
& - \overset{k}{\underset{h=1}{S}} A_h^{r_2} B_h^{r_4} \overset{k}{\underset{h=1}{S}} \left[(-n_1)^{\alpha_h+\beta_h} - 1\right] A_h^{r_1} B_h^{r_3} - \overset{k}{\underset{h=1}{S}} A_h^{r_1} B_h^{r_3} \\
& \overset{k}{\underset{h=1}{S}} \left[(-n_1)^{\alpha_h'+\beta_h'} - 1\right] A_h^{r_2} B_h^{r_4} + Q_{r_1 r_3} Q_{r_2 r_4} - Q_{(r_1+r_2)(r_3+r_4)}. \qquad (3.38)
\end{aligned}$$

Finally, combining the various equations we get the formula:

$$\begin{aligned}
{}_2M_{p_{ab}} = {} & n^{-1}\bar{p}_{2a2b} - \bar{p}^2_{a\,b} + 2(n)^{-2(a+b+1)}\,(a!)^2(b!)^2 \overset{n}{\underset{i_h=1}{S}} \overset{a,b}{\underset{\alpha_h,\alpha_h',\beta_h,\beta_h'=0}{S}} \overset{t}{\underset{k=1}{S}} \\
& \left\{ \overset{k}{\underset{h,h'=1}{S}} (-n_1)^{\alpha_h+\beta_h+\alpha_{h'}'+\beta_{h'}'} + n_k \overset{k}{\underset{h=1}{S}} [(-n_1)^{\alpha_h+\beta_h} + (-n_1)^{\alpha_h'+\beta_h'}] + C_2^{nk} \right\} \\
& \frac{\Pi P^{ih}\,(\alpha_h+\alpha_h')\,(\beta_h+\beta_h')}{\Pi\alpha_h!\,\Pi\beta_h!\,\Pi\alpha_h'!\,\Pi\beta_h'!} + 2(n)^{-2(a+b+1)}\,(a!)^2(b!)^2 \overset{n}{\underset{i_h=1}{S}} \overset{a,b}{\underset{r_1,r_2,r_3,r_4=0}{S}} \frac{(-n)^{Sr_1} Sr_i}{\Pi r_1!} \\
& \overset{\alpha,\beta,\gamma,\delta}{\underset{\alpha_h,\alpha_h',\beta_h,\beta_h'=0}{S}} \overset{t}{\underset{k=1}{S}} \left\{ \overset{k}{\underset{h,h'=1}{S}} [(-n_1)^{\alpha_h+\beta_h+\alpha_{h'}'+\beta_{h'}'} - 1]\, A_h^{r_1} B_h^{r_3} A_{h'}^{r_2} B_{h'}^{r_4} \right. \\
& - \overset{k}{\underset{h=1}{S}} [(-n_1)^{\alpha_h+\beta_h} - 1]\, A_h^{r_1} B_h^{r_3} \overset{k}{\underset{h=1}{S}} A_h^{r_2} B_h^{r_4} - \overset{k}{\underset{h=1}{S}} [(-n_1)^{\alpha_h'+\beta_h'} - 1] \\
& A_h^{r_2} B_h^{r_4} \overset{k}{\underset{h=1}{S}} A_h^{r_1} B_h^{r_3} + Q_{r_2 r_4} \overset{k}{\underset{h=1}{S}} [(-n_1)^{\alpha_h+\beta_h} - 1]\, A_h^{r_1} B_h^{r_3}
\end{aligned}$$

$$+ Q_{r1r3} \sum_{h=1}^{k} [(-n_1)^{\alpha_h'+\beta_h'} - 1] A_h^{r_2} B_h^{r_4} + Q_{r1r3} Q_{r2r4} - Q_{(r1+r2)\,(r3+r4)} \Bigg\}$$

$$\frac{\Pi P^{j_h}_{(\alpha+\alpha_h')\,(\beta_h+\beta_h')}}{\Pi\alpha_h!\,\Pi\beta_h!\,\Pi\alpha_h'!\,\Pi\beta_h'!}. \tag{12}$$

In case the n populations are identical the second part of (12) must vanish, and in the first part the summations

$$\sum_{j_h=1}^{n} \prod_{h=1}^{k} P^{j_h}_{(\alpha_h+\alpha_h')\,(\beta_h+\beta_h')} = \frac{k!\, C_k^n \,\Pi\, P_{(\alpha_h+\alpha_h')\,(\beta_h+\beta_h')}}{l_1!\, l_2! \cdots l_c!},$$

where $l_1, l_2, \cdots l_c$ are the number of repetitions of the pairs of integers $(\alpha_1 + \alpha_1')\,(\beta_1 + \beta_1'), \cdots (\alpha_k + \alpha_k')\,(\beta_k + \beta_k')$, respectively.

We then have the following

Corollary: The mathematical expectation of the variance, ${}_2m_{pab}$, of the product moment, p_{ab}, in samples of n from a single infinite population is given by

$${}_2M_{pab} = \bar{p}_{2a2b} - \bar{p}_{ab}^2 + 2(n)^{-2(a+b+1)}\,(a!)^2(b!)^2 \sum_{\alpha_h,\alpha_{h'}',\beta_h,\beta_{h'}'=0}^{a,a,b,b}$$

$$\sum_{k=1}^{t} \frac{k!\, C^{nk}}{l_1!\, l_2! \cdots l_c!} \Bigg\{ \sum_{h,h'=1}^{k} (-n_1)^{\alpha_h+\beta_h+\alpha_{h'}'+\beta_{h'}'} + n_k \sum_{h=1}^{k} [(-n_1)^{\alpha_h+\beta_h}$$

$$+ (-n_1)^{\alpha_h'+\beta_h'}] + C_2^{nk} \Bigg\} \frac{\prod_1^k P(\alpha_h + \alpha_h')\,(\beta_h + \beta_h')}{\Pi\alpha_h!\,\Pi\beta_h!\,\Pi\alpha_h'!\,\Pi\beta_h'!}. \tag{12'}$$

4. **The Formula for ${}_2M_{p21}$.** Formula (12) can by no means be used mechanically. It does, however, summarize to a great extent the details in finding ${}_2M_{pab}$ for any given values a, b. Formulae for ${}_2M_{p21}$, ${}_2M_{p31}$ have been obtained, but the one for ${}_2M_{p31}$ is too long to be included in the paper, especially since with a little work it can be easily derived by applying (12). The one for ${}_2M_{p21}$ is given immediately below.

$${}_2M_{p21} = n^{-6}\{n_1^2 n_2^2 S[P^i_{42} - (P^i_{21})^2] + n_2^2 S[P^i_{40}P^i_{02} + 4(P^i_{30}P^j_{12} - n_2 P^i_{31}P^j_{11})]$$

$$- 2n_2^2 n_3 S P^i_{22}P^j_{20} + (n_2^2 + 2)S(P^i_{20}P^j_{20}P^k_{02} + 8P^i_{20}P^j_{11}P^k_{11}) + 6SP^i_{20}P^j_{20}P^k_{02}\}$$

$$+ 2n^{-5}\{n_1 n_2^2 S(P^i_{41}B_i + 2P^i_{32}A_i - P^i_{30}P^i_{11}B_i - 2P^i_{21}P^i_{11}A_i)$$

$$- 4n_2 n_4 S(P^i_{30}B_iP^j_{11} + P^i_{12}A_iP^j_{20}) - 2n_2 S[n_3 P^i_{21}B_iP^j_{20} + 2(2n_2 - 3)P^i_{21}A_iP^j_{11}]$$

$$+ 6nSP^i_{21}P^j_{11}A_j + 4n_2 S(P^i_{30}P^j_{11}B_j + P^i_{30}P^j_{02}A_j + P^i_{30}A_iP^j_{02} + 2P^j_{21}P^j_{20}B_j$$

$$+ P^i_{12}P^j_{20}A_j)\} + n^{-4}\{n_2^2 S[P^i_{40}B_i^2 - (P^i_{20}B_i)^2] + 4SP^i_{20}P^j_{20}(B_i + B_j)^2$$

$$+ 3(n_2^2 + n_2)SP^i_{22}A_i^2 + 4SP^i_{20}P^j_{02}(A_i + A_j)^2 - 2n_4 S[P^i_{20}P^j_{02}A_j^2$$

$$+ 2P^i_{11}P^j_{11}(A_i + A_j)^2] + 16SP^i_{11}A_iP^j_{11}A_j - 4n_2^2S(P^i_{11}A_i)^2$$
$$+ 4(2n_2^2 + n_2)SP^i_{31}A_iB_i - 4n_4SP^i_{11}A_iB_iP^j_{20} - 8n_3SP_{11}P^j_{20}A_jB_j$$
$$+ 8S(P^i_{11}B_iP^i_{20}A_j + P^i_{11}A_iP^j_{20}B_j) - 4n_2^2SP^i_{11}A_iP^i_{20}B_i$$
$$- 2n_1n_2n^{-1}S(Q_{20}P^i_{22} + 2Q_{11}P^i_{31}) + 2n_2n^{-1}S[6Q_{11}P^i_{11}P^j_{20}$$
$$+ Q_{20}(P^i_{20}P^j_{02} + 4P^i_{11}P^j_{11})]\} + 2n^{-4}\{nn_2S(2P^i_{30}A_iB_i^2 + 2P^i_{12}A_i^3 + 5P^i_{21}A_i^2B_i)$$
$$- n_1S[Q_{20}(P^i_{21}B_i + P^i_{12}A_i) + 2Q_{11}8P^i_{30}B_i + 2P^i_{21}A_i)]\} + n^{-4}\{n^2S[P^i_{02}A_i^4$$
$$+ 4(P^i_{20}A_i^2B_i^2 + P^i_{11}A_i^2B_i)] - 2nS[(Q_{20}A_iB_i + Q_{11}A_i^2)P^i_{11} + Q_{11}P^i_{20}A_iB_i$$
$$+ Q_{20}P^i_{02}A_i^2] + S[Q^2_{20}P^i_{02} + 4Q_{20}(Q_{11}P^i_{11} + Q^2_{11}P^i_{20})]\}. \qquad (13)^{12}$$

CHAPTER IV. The Mathematical Expectation of the Third Moment of p_{11}

1. The Mathematical Expectation of ${}_3m_{p_{11}}$. Following the notation of the last chapter we shall denote the third moment of p_{11} about its mean by ${}_3m_{p_{11}}$ and the mathematical expectation of ${}_3m_{p_{11}}$ by ${}_3M_{p_{11}}$. We have then by definition.

$${}_3m_{p_{11}} = \{n^{-1}S(x_i - x)(y_i - y) - \bar{p}_{11}\}^3,$$

and by a well known formula we have:

$${}_3M_{p_{11}} = \overline{p^3_{11}} - 3_2M_{p_{11}}\bar{p}_{11} - \bar{p}^3_{11}. \qquad (4.11)$$

The last two terms of (4.11) are given by (1) and (11). To evaluate $\overline{p^3_{11}}$ we write:

$$\overline{\bar{p}^3_{11}} = E\{n^{-1}S(x_i - x)(y_i - y)\}^3 = n^{-3}SE(x_i - x)^3(y_i - y)^3$$
$$+ 3n^{-3}SE(x_i - x)^2(y_i - y)^2(x_j - x)(y_j - y)$$
$$+ 6n^{-3}SE(x_i - x)(y_i - y)(x_j - x)(y_j - y)(x_k - x)(y_k - y).$$

The first term is simply $n^{-2}\bar{p}_{33}$ which is given by (10). The evaluation of the second term is not essentially different from the evaluation of the left hand side of (3.22), and since all details have been given there we shall omit them here.

To evaluate the last expression let us write:

$$SE(x_i - x)(y_i - y)(x_j - x)(y_j - y)(x_k - x)(y_k - y)$$
$$= SE[(U_i + A_i)(V_i + B_i)(U_j + A_j)(V_j + B_j)(U_k + A_k)(V_k + B_k)]$$
$$= SE(U_iV_iU_jV_jU_kV_k) + SE(U_iV_iU_jV_jU_kB_k) + \cdots + SE(A_iB_iA_jB_jA_kB_k). \qquad (4.12)$$

---

[12] In case the n populations are identical this reduces to one of Pepper's formulae, *Biometrika*, Vol. XXI, p. 238, Cor. 1.

As there is a great deal of similarity among the various terms of the right hand side of (4.12), it will not be necessary to go into the details of the expansion of every one of them. We shall, therefore, indicate the details for the expansion of only two of them—one symmetrical and one non-symmetrical; and as the first two terms are of that type we shall use these for the purpose of illustration.

Using the u, v notation we have

$$SE(U_iV_iU_jV_jU_kV_k) = n^{-6}SE[(n_1u_i - \cdots)(n_1v_i - \cdots)(n_1u_j - \cdots)(n_1v_j - \cdots)(n_1u_k - \cdots)(n_1u_k - \cdots)].$$

The maximum number of subscripts appearing in any term evidently being 3, we can write without any loss in generality:

$$SE[(n_1u_i - \cdots)\cdots(n_1v_k - \cdots)] = E[(n_1u_1 - \cdots)(n_1v_1 - \cdots)(n_1u_2 - \cdots)$$
$$n_1v_2 - \cdots)(n_1u_3 - \cdots)(n_1v_3 - \cdots)] + E\{(n_1u_1 - \cdots)(n_1v_1 - \cdots)[(n_1u_2 - \cdots)$$
$$(n_1v_2 - \cdots) + (n_1u_3 - \cdots)(n_1v_3 - \cdots)] + (n_1u_2 - \cdots)(n_1v_2 - \cdots)$$
$$(n_1u_3 - \cdots)(n_1v_3 - \cdots)\}\underset{4}{S}(n_1u_i - \cdots)(n_1v_i - \cdots) + E\{(n_1u_1 - \cdots)$$
$$(n_1v_1 - \cdots) + (n_1u_2 - \cdots)(n_1v_2 - \cdots) + n_1u_3 - \cdots)(n_1v_3 - \cdots)\}$$
$$\underset{4}{S}(n_1u_i - \cdots)(n_1v_i - \cdots)(n_1u_j - \cdots)(n_1v_j - \cdots) + \underset{4}{SE}\{(n_1u_i - \cdots)\cdots$$
$$(n_1v_k - \cdots)\}. \qquad (4.13)$$

The coefficients of the various terms arising in this expansion can now be found quite easily. For example, the coefficient of $P^1_{3\,3}$, which is, of course, the same as the coefficient of $P^i_{3\,3}$, is easily found to be

$$n_1^2 + n_3(2n_1^2 + 1) + \frac{n_3n_4(n_1^2 + 2)}{2} + \frac{n_3n_4n_5}{6} = \frac{nn_1n_2(3n_1 - 2)}{6}.$$

To evaluate the summation $SE(U_iV_iU_jV_jU_kB_k) = n^{-5}SE[(n_1u_i - \cdots)(n_1v_i - \cdots)(n_1u_j - \cdots)(n_1v_j - \cdots)(n_1v_k - \cdots)B_k]$, we break it up into partial summations as follows:

$$SE[(n_1u_i - \cdots)(n_1v_i - \cdots)(n_1u_j - \cdots)(n_1v_j - \cdots)(n_1u_k - \cdots)B_k]$$
$$= E\{(n_1u_1 - \cdots)(n_1v_1 - \cdots)[(n_1u_2 - \cdots)(n_1v_2 - \cdots)(n_1u_3 - \cdots)B_3$$
$$+ (n_1u_2 - \cdots)B_2(n_1u_3 - \cdots)(n_1v_3 - \cdots)] + (n_1u_1 - \cdots)B_1(n_1u_2 - \cdots)$$
$$(n_1v_2 - \cdots)(n_1u_3 - \cdots)(n_1v_3 - \cdots)\} + E\{(n_1u_1 - \cdots)(n_1v_1 - \cdots)$$
$$[(n_1u_2 - \cdots)(n_1v_2 - \cdots) + (n_1u_3 - \cdots)(n_1v_3 - \cdots)] + (n_1u_2 - \cdots)$$
$$(n_1v_2 - \cdots)(n_1u_3 - \cdots)(n_1v_3 - \cdots)\}\underset{4}{S}(n_1u_j - \cdots)B_j + E\{(n_1u_1 - \cdots)$$
$$(n_1v_1 - \cdots)[(n_1u_2 - \cdots)B_2 + (n_1u_3 - \cdots)B_3] + (n_1u_2 - \cdots)(n_1v_2 - \cdots)$$

$$[(n_1u_1 - \cdots)B_1 + (n_1u_3 - \cdots)B_3] + (n_1u_3 - \cdots)(n_1v_3 - \cdots)$$
$$[(n_1u_1 - \cdots)B_1 + (n_1v_2 - \cdots)B_2]\}\underset{4}{S}(n_1u_j - \cdots)(n_1v_j - \cdots)$$
$$+ E\{(n_1u_1 - \cdots)(n_1v_1 - \cdots) + (n_1u_2 - \cdots)(n_1v_2 - \cdots) + (n_1u_3 - \cdots)$$
$$(n_1v_3 - \cdots)\}\underset{4}{S}(n_1u_i - \cdots)(n_1v_i - \cdots)(n_1u_j - \cdots)B_j + E\{(n_1u_1 - \cdots)B_1$$
$$+ (n_2u_2 - \cdots)B_2 + (n_1u_3 - \cdots)B_3\}\underset{4}{S}(n_1u_i - \cdots)(n_1v_i - \cdots)$$
$$(n_1u_j - \cdots)(n_1v_j - \cdots) + E\underset{4}{S}(n_1u_i - \cdots)(n_1v_i - \cdots)(n_1u_j - \cdots)$$
$$(n_1v_j - \cdots)(n_1u_k - \cdots)B_k. \tag{4.14}$$

The expansion of (4.14) is not as difficult as it appears for only two subscripts can appear in any term: the explicit appearance of the subscript 3 is due to the fact that we are dealing with a triple summation. We, consequently, do not need to expand those parentheses in which B appears.

We shall now, without any further details, state the final result, which is:

$$\begin{aligned}
{}_3M_{p_{11}} &= n^{-6}\{S[n_1^3P^i_{33} - P^i_{30}P^j_{03} + 3n_1(P^i_{31}P^j_{02} + P^i_{20}P^j_{13}) + 3n_1(n_1^2+2)P^i_{22}P^j_{11} \\
&\quad - 3(2n_1^2+1)P^i_{21}P^j_{12} + 3n_3P^i_{11}P^j_{02}P^k_{20} + 6(n_1^3 + 3n_1 - 2)P^i_{11}P^j_{11}P^k_{11}] \\
&\quad - 3n_1SP^i_{11}[S(n_1^2P^i_{22} + P^i_{20}P^j_{02} - n_1^2(P^i_{11})^2 + 2P^i_{11}P^j_{11})] - n_1^3(SP^i_{11})^3]\} \\
&\quad + 3n^{-5}\{S[n_1^2(P^i_{32}B_i + P^i_{23}A_i) + 2\alpha(P^i_{21}P^j_{11}B_i + P^i_{12}P^j_{11}A_i) \\
&\quad - 2n_1(P^i_{21}P^i_{11}B_i + P^i_{12}P^i_{11}A_i) - 2n_1(P^i_{11}P^j_{21}B_i + P^i_{11}P^j_{12}A_i) \\
&\quad + (P^i_{12}P^j_{20}B_i + P^i_{21}P^j_{02}A_i) - 2n_1(P^i_{12}P^j_{20}B_j + P^i_{21}P^j_{02}A_j) \\
&\quad + (P^i_{30}P^j_{02}B_j + P^i_{03}P^j_{20}A_j)]\} + 3n^{-4}\{S[n_1(P^i_{31}B_i^2 + P^i_{13}A_i^2) \\
&\quad + n_2(P^i_{20}P^i_{11}B_i^2 + P^i_{02}P^j_{11}A_i^2) - (P^i_{11}P^i_{20}B_i^2 + P^i_{02}P^i_{11}A_i^2) \\
&\quad - 2(P^i_{20}B_iP^j_{11}B_j + P^i_{02}A_iP^j_{11}A_j) + 2n_1P^i_{22}A_iB_i - 2P^i_{20}B_iP^j_{02}A_j \\
&\quad - 2P^i_{11}A_iP^j_{11}B_j + 2n_2P^i_{11}P^j_{11}A_iB_i - 2(P^i_{11})^2A_iB_i]\} + n^{-3}\{S[(P^i_{30}B_i^3 + P^i_{03}A_i^3) \\
&\quad + 3(P^i_{21}A_iB_i^2 + P^i_{12}A_i^2B_i)]\}.
\end{aligned} \tag{14}^{13}$$

Where $\alpha = n_1^2 + n_1 + 1$.

This formula is shorter and simpler than the formula for ${}_2M_{p_{21}}$, although they are of the same order. This is due to the symmetry of ${}_3M_{p_{11}}$.

Chapter V. Product Moments of Trivariate and Quadrivariate Populations

1. Some additional definitions and notation. In this chapter we shall indicate briefly how the method of the previous chapters may be extended to populations

[13] Cf. *Biometrika*, Vol. XXI, p. 253, formula (19).

of more than two variables. We shall do this by deriving some of the simpler formulae, corresponding to those of Chapter II, for trivariate and quadrivariate populations.

The notation will be slightly changed in that we shall symbolize the new variables by priming the symbols for the variables used in the previous chapters. Thus, we shall indicate the k^{th} trivariate population by (X_k, Y_k, X'_k) and the k^{th} quadrivariate population by (X_k, Y_k, X'_k, Y'_k), and samples from such populations by (x_k, y_k, x'_k) and (x_k, y_k, x'_k, y'_k) respectively.

We shall denote by P^m_{ijk} the product moment of the m^{th} population of order i in X, j in Y, and k in X', and by P^m_{ijkl} the similar product moment for a quadrivariate population. These are defined by the following equations:

$$P^m_{ijk} = E(X_m - a_m)^i(Y_m - b_m)^j(X'_m - c_m)^k, \tag{5.11}$$

$$P^m_{ijkl} = E(X_m - a_m)^i(Y_m - b_m)^j(X'_m - c_m)^k(Y_m - d_m)^l \tag{5.12}$$

where a_m, b_m, etc. are defined as in Chapter I part 2.

The sample product moments corresponding to P^m_{ijk}, P^m_{ijkl} will be denoted by p_{ijk} and p_{ijkl} respectively. They are defined by:

$$p_{ijk} = n^{-1} \sum_{m=1}^{n} (x_m - x)^i(y_m - y)^j(x'_m - x')^k, \tag{5.13}$$

$$p_{ijkl} = n^{-1} \sum_{m=1}^{n} (x_m - x)^i(y_m - y)^j(x'_m - x')^k(y'_m - y')^l. \tag{5.14}$$

Finally we shall designate $E(p_{ijk})$ and $E(p_{ijkl})$ by $\bar{p}_{ijk}$ and $\bar{p}_{ijkl}$ respectively.

2. **The Mathematical Expectation of p_{111} and p_{211}.** By definition we have

$$\bar{p}_{111} = E[n^{-1}S(x_i - x)(y_i - y)(x'_i - x')]. \tag{5.21}$$

Applying the transformations (1.17) this equation becomes

$$np_{111} = E[S(U_i + A_i)(V_i + B_i)(U'_i + C_i)] = SE(U_iV_iU'_i) + SE(U_iV_iC_i)$$
$$+ SE(U_iU'_iB_i) + SE(V_iU'_iA_i) + \text{vanishing terms} + SE(A_iB_iC_i). \tag{5.22}$$

Since $EA_iB_iC_i = A_iB_iC_i$, $SE(A_iB_iC_i) = SA_iB_iC_i$. Following the previous notation we shall put $SA_iB_iC_i = Q_{111}$.

When the expression $SE(U_iV_iU'_i)$ is expanded, no other non-vanishing terms except those of the form $E(u_iv_iu'_i) = P^i_{111}$ can appear. The coefficient of this term will evidently be the same as that of P^i_{21} in (2.23), namely: $n^{-2}n_1n_2$. Whence:

$$SE(U_iV_iU'_i) = n^{-2}n_1n_2SP^i_{11}.$$

The three terms following the first of (5.22) are by (2.24) equal to

$$n^{-1}n_2S(P^i_{110}C_i + P^i_{101}B_i + P^i_{011}A_i).$$